A PROSPECTIVE CLINICO-RADIOLOGICAL STUDY OF PROGRESSIVE SUPRANUCLEAR PALSY USING SERIAL MRI WITH REGISTRATION

Submitted to the University of London

PhD Thesis

Dominic Curtis Paviour

Institute of Neurology, University College London

2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The purpose of this thesis is to investigate: whether cross sectional MRI can help to discriminate progressive supranuclear palsy (PSP), multiple system atrophy (Parkinsonian subtype, MSA-P) and Parkinson's disease (PD); whether established methods of serial volumetric MRI analysis can be applied to measure in-vivo rates of brain atrophy in PSP and; whether cross sectional MRI and MRI-derived atrophy rates have a clinical correlate.

Volumetric and diffusion-weighted MRI (DWI) scans were performed in 24 patients with PSP, 11 with MSA-P, 12 with PD and 18 healthy controls. Detailed clinical and neuropsychological assessments were undertaken, and results at the time of initial assessment and follow-up compared.

Whole brain and regional brain volumes were measured on positionally matched (registered) baseline and follow up MR images. The boundary shift integral (BSI), an image analysis technique for assessing volume differences from registered MR scans was applied to whole brain, and hypothesis-driven regions of interest in order to allow quantification of atrophy rates.

- PSP can be reliably discriminated from MSA-P with quantitative regional volume measurements of the superior cerebellar peduncle, midbrain, pons and cerebellum ($p<0.001$).
- DWI may help to identify regional pathological change *in-vivo*, discriminating MSA-P and PSP using apparent diffusion coefficients in the middle cerebellar peduncle ($p<0.001$).
- The mean(SD) brain atrophy rate in PSP, 1.2 (1.0) $\%year^{-1}$ was greater than in healthy controls 0.4(0.5)$\%year^{-1}$ ($p=0.04$) but similar to MSA-P 1.0(1.1)$\%year^{-1}$. In PSP, the

mean midbrain atrophy rate was most significantly different from controls ($p<0.001$). Ponto-cerebellar atrophy rates discriminated MSA-P and PSP ($p<0.05$).

The distinct patterns and rates of atrophy in these diseases have a meaningful clinical correlate and power calculations confirm that in PSP, regional atrophy rates are quantifiable and feasible markers of disease progression that have utility in clinical trials as a marker for evaluating disease-modifying treatments in PSP.

Table of Contents

Tables

Figures

Figures 3.4, 3.5, and 10.2-10.7 are included on a CD-Rom.

Abbreviations Used Throughout

6dof/9dof	Six/nine degrees of freedom
AD	Alzheimer's disease
ADC	Apparent diffusion coefficient
ADL	Activities of daily living
ANOVA	Analysis of variance
BBSI	Brain boundary shift integral
CBD	Cortico-basal degeneration
CNS	Central nervous system
CSF	Cerebrospinal fluid
CT	Computed tomography
DBC	Differential bias correction
DWI	Diffusion weighted MR imaging
FAB	Frontal assessment battery
FDR	False discovery rate
FEF	Frontal eye fields
FTD	Frontotemporal dementia
FTDP-17q	Frontotemporal dementia with Parkinsonism linked to Chromosome 17
FTLD-MND	Frontotemporal lobar degeneration with MND
FTLD-U	Frontotemporal lobar degeneration with Ubiquitin Inclusions
GP	Globus pallidus (GPi) Internus
H+Y	Hoehn and Yahr rating scale
ICC	Intraclass correlation coefficient
Mattis DRS	Mattis Dementia Rating scale
MIDAS	Medical Image Display and Analysis Software
MMSE	Mini mental state examination
MND	Motor neuron disease
MRI	Magnetic resonance imaging
MSA-P	Multiple system atrophy (Parkinsonian subtype)
NFT	Neurofibrilliary tangles
NINDS-SPSP	National Institute for Neurological diseases and Stroke- Society for PSP
NRI	Nucleus raphe interpositus
NT	Neuropil threads
PASAT	Paced auditory serial additions test

PD	Parkinson's disease
PDSBB	Parkinson's disease society brain bank
PSP	Progressive supranuclear palsy
Rey AVLT	Rey auditory verbal learning test
RF	radio frequency
RMT	Recognition memory test for faces
RTMT A/B	Reitan trail making test A/B
SCP	superior cerebellar peduncle
SD	Standard deviation
SN	Substantia nigra
SNR	Signal to noise ratio
SPM	Statistical parametric mapping
STN	Subthalamic nucleus
T	Tesla
TE	Time to echo
TR	Time to repetition
UPDRS	Unified Parkinson's disease rating scale
VBM	Voxel based morphometry
WAIS-R	Weschler Adult Intelligence scale-revised
WCST	Wisconsin card sorting test

Outline

Progressive supranuclear palsy (PSP) is the commonest of the neurodegenerative extrapyramidal syndromes after Parkinson's disease (PD). Patients present with an insidious onset of neurological symptoms and signs including impaired balance, non-specific visual symptoms, impaired mobility and cognitive disturbance. Even with the classical presentation, diagnosis is difficult and atypical presentations with Parkinsonism and gait freezing complicate matters further. There are no biomarkers or pathognomonic neuroimaging features and a definitive diagnosis requires histopathological confirmation, usually at autopsy. The clinical diagnosis can only be reliably made when typical clinical signs are present, and this may not be the case, particularly in the early stages of the disease.

Magnetic resonance imaging (MRI) of the brain at a single time point has for some time been a standard investigation and an obviously small midbrain, is supportive of the clinical diagnosis of PSP. Other pathologically affected regions may have been overlooked and newer imaging and image analysis techniques, now more widely available, may soon prove to be helpful in discriminating the degenerative bradykinetic rigid disorders.

There are no efficacious disease modifying drugs and PSP progresses relentlessly with a mean duration to death of between six and nine years. The cause is obscure but the disease is thought to result from a combination of environmental and genetic factors. Abnormal accumulation of hyper-phosphorylated Tau aggregates is now known to be an important pathological process and this understanding may pave the way for new therapeutic approaches. Methods for assessing disease progression in PSP are lacking. Outcome measures in clinical trials include clinical measures of disease progression, which by their nature involve subjective assessment of disease severity, are affected

by floor and ceiling effects and may be non-linear. A subject's ability to perform on these tests may fluctuate on a day-to-day basis and symptomatic effects of a drug may confuse the issue by improving test scores. Some of these problems can be overcome by adapting trial design appropriately, but clinical rating scores alone may still not provide truly sensitive and reliable markers of disease progression.

In PSP, the inevitable downstream effect of abnormal protein deposition and cell death is accelerated cerebral atrophy. This needs to be distinguished from the slower loss of neurones seen as part of the normal ageing process. In PSP, whole brain atrophy may be a good surrogate marker of disease progression. However because the most severe macroscopic atrophy at post mortem is seen in the brainstem, regional atrophy rates may provide a better marker of disease progression.

Imaging at a single time point can only infer that atrophy has occurred. In order to quantify true rates of brain atrophy, it is necessary to compare serial scans from the same individual over time. To facilitate comparison, these scans can be digitally matched (registered) to one another with a high degree of precision. Rates of global and regional atrophy can then be calculated, by comparing brain volumes over time, by assessing expansion of cerebrospinal fluid (CSF) spaces or by using an automated method of assessing brain volume change over time such as the brain boundary shift integral (BBSI). Furthermore, a local boundary shift integral can be applied to limit automatic calculation of atrophy to a predetermined region.

The research undertaken for this thesis was designed to determine rates of brain atrophy, derived from serial MRI scans, in PSP compared to established rates of atrophy in healthy controls and other bradykinetic rigid neurodegenerative conditions. Specifically, the aims were: (1) to look for new MRI based diagnostic markers in PSP; (2) to use novel image analysis techniques to determine the patterns of progressive

global brain atrophy occurring in PSP, multiple system atrophy of the Parkinsonian subtype (MSA-P) and Parkinson's disease; (3) to compare rates of whole brain atrophy in PSP, MSA-P, PD and healthy controls; (4) to compare regional rates of brain atrophy in PSP, MSA-P, PD and healthy controls and to determine if these provided better markers of disease progression; and (5) to look for correlations between clinical features, clinical rating scales, neuropsychological scores and MRI findings.

As a consequence of these aims, I hoped to determine whether MRI could be applied as a method of measuring brain atrophy and disease progression in PSP, as well as exploring its potential as an ancillary aid in the differential diagnosis of the bradykinetic rigid disorders.

1. Introduction

1.1 Degenerative bradykinetic rigid syndromes

In progressive supranuclear palsy (PSP), multiple system atrophy (MSA) and Parkinson's disease (PD), bradykinesia and rigidity occur as a result of degeneration of basal ganglia structures including the globus pallidus, the substantia nigra, the corpus striatum and the subthalamic nucleus. Many other brain regions are damaged in PSP and MSA resulting in the additional clinical features characteristic of these nosological entities.

1.2 Progressive supranuclear palsy

1.2.1. Historical perspective

Progressive supranuclear palsy (PSP) was first described in the medical literature in 1963 (Richardson *et al.*, 1963) and in more detail a year later in (Steele *et al.* 1964). Nine cases with a distinct clinical syndrome consisting of, initially vague visual or speech difficulties, changes in personality and an unsteady gait were described. A supranuclear ophthalmoplegia was a constant and distinctive finding, occurring early in the disease course. Barbeau called the disorder "Steele-Richardson-Olszewski syndrome", after the three authors of these initial reports (Barbeau, 1965), and this rubric is still used in some parts of the world.

The Canadian authors believed this clinical picture represented a distinct disease entity and they conceded that it was unlikely the disease they described was a new one, but felt it had probably been mislabelled as "arteriosclerotic" Parkinsonism. They expressed a degree of surprise that for a syndrome with such a characteristic clinical

appearance, there was no detailed clinico-pathological description in the medical literature.

Fragmentory descriptions, however can be found in the early medical literature and Charles Dickens in "The Lazy Tour of Two Idle Apprentices" writes of *"A chilled, slow, earthy, fixed old man. A cadaverous old man of measured speech. An old man who seemed as unable to wink, as if his eyelids had been nailed to his forehead. An old man whose eyes - two spots of fire - had no more motion than if they had been connected with the back of his skull by screws driven through it, and riveted and bolted outside, among his grey hair..... He had come in and shut the door, and he now sat down. He did not bend himself to sit, as other people do, but seemed to sink bolt upright, as if in water, until the chair stopped him"* (Dickens C, 1990).

Dickens seems to describe quite clearly: the paucity of movement or "body bradykinesia"; the slow, growling dysarthria; the frontalis over-activity and staring appearance; the gaze palsy; and the tendency to sit "en-bloc" (Larner, 2002).

Although it is not possible with certainty to know whether the individuals described in the novel or the early medical patient descriptions had PSP, it is unlikely that PSP is a new disease.

1.2.2. Epidemiology

The geographical distribution of PSP is not well described and it is presumed that in the western world, the population prevalence is similar from country to country. Certain geographic isolates have reported unexpectedly high rates of atypical Parkinsonism, with a clinical presentation similar to PSP, contrasting with that seen in Europe and North America. The island of Guadeloupe in the French West Indies is

one such area. The characteristic clinical picture of Guadeloupean atypical parkinsonism is of symmetrical bradykinesia, axial rigidity, early postural instability, cognitive decline with prominent frontal lobe dysfunction and a negligible response to levodopa (Caparros-Lefebvre *et al.*, 2002). It was thought that the disease was related to the ethnicity of the population or to the consumption of herbal teas containing alkaloid toxins. The most recent findings suggest that it is a tauopathy with some unusual pathological features.

A high incidence of PSP has been described in New Caledonia (Angibaud *et al.*, 2004) and the occurrence of a high incidence of parkinsonism-dementia complex (PDC) among Chamorros on Guam was also first described fully in the 1960's (Hirano *et al.*, 1961a; Hirano *et al.*, 1961b) and it is recognised as a distinct disease entity on the basis of clinico-pathological studies. However it has many features which closely resemble PSP including a supranuclear gaze palsy in some cases. More recently, a first pathologically confirmed case of PSP has been reported on the island of Guam.

Large studies of population prevalence are prone to under-ascertainment of cases and small community studies, while thorough, are only able to identify a small number of cases. Ten years ago, the prevalence of PSP may have been quoted at around 1 in 100,000 but PSP is more common than previously considered and recently, the prevalence has been adjusted to 5 per 100,000, with an age adjusted prevalence of 6.4 per 100,000, based on detailed epidemiological studies (Nath *et al.*, 2001; Schrag *et al.*, 1999). This means it is about as common as motor neuron disease, a condition with a much greater level of public and professional awareness. Nath *et al* undertook national, regional and community prevalence studies. The national studies relied upon

the British Neurological Surveillance unit, the PSP (Europe) Association and the Office of National Statistics for case identification. In addition, neurologists from across the UK were invited to refer cases. In the regional study, direct referral of cases, a review of 33,000 out patient letters, screening of regional databases such as gastrostomy tube registers and movement disorder registers as well as in-patient diagnosis coding registers were reviewed. In the community study, general practice case records were screened for cases that might fit with a diagnosis of PSP and these were invited to clinic or examined at home. Cases were included if they conformed to the NINDS-SPSP diagnostic categories for possible or probable PSP (Litvan *et al.*, 1996a). It was presumed that the national study would significantly underestimate the prevalence of PSP, which it did with a crude prevalence of 1 per 100,000. The regional study identified an age-adjusted prevalence of 2.4 per 100,000 and the community study identified a prevalence of 5 per 100,000. Clearly the methods of case finding and the size of the surveyed population affect the estimated prevalence but overall, the clear message is that PSP is under recognised. A lack of awareness of the condition amongst physicians and phenotypic variability may explain the diagnostic underascertainment.

1.2.3. Clinical features and differential diagnosis of PSP

Gait instability, impairment of vertical eye movements, spastic or ataxic dysarthria, dysphagia, bradykinesia, rigidity, frontal behavioural changes and sub-cortical dementia are the cardinal signs of PSP. Most cases begin after sixty years of age and result in death six to nine years after disease onset (Lees, 1987).

Patients with typical PSP tend to present with gait disturbance and unsteadiness as well as a marked tendency to topple backwards. The gait has a characteristic reeling and staggering quality due to the stiff posture of the trunk and the neck and is sometimes likened to that of a "dancing bear". Whilst a supranuclear gaze palsy with downgaze abnormalities is highly characteristic, it is often not present in the early stages and early visual symptoms may be non-specific and difficult to diagnose. Photophobia, watering of the eyes, blurred vision, difficulty focussing on and following written text have all been reported in the early stages of the disease, prior to the clinical diagnosis being made when other more definitive clinical symptoms and signs emerge.

Various clinical diagnostic criteria for PSP have been proposed over the years (Lees, 1987) (Blin *et al.*, 1990; Collins *et al.*, 1995; Duvoisin RC., 1992; Golbe LI, 1993; Tolosa *et al.*, 1994). The NINDS-SPSP diagnostic criteria published in 1996 (Litvan *et al.*, 1996a) and revised in 2003 (Litvan *et al.*, 2003) is the most commonly used for research purposes and are outlined below (see also Chapter 3, table 3.6a):

For possible and probable PSP: A gradually progressive disorder with age at onset at 40 years or later.

Possible PSP: Either a vertical supranuclear palsy, or both slowing of vertical saccades and postural instability with falls less than 1 year after disease onset.

Probable PSP: A vertical supranuclear palsy and prominent postural instability with falls within first year of disease onset. Later defined as falls or the tendency to fall (patients are able to stabilize themselves).

Definite PSP: All criteria for possible or probable PSP are met and there is histopathological confirmation at autopsy.

Supportive features: Symmetrical akinesia or rigidity, proximal more than distal; abnormal neck posture, especially retrocollis; poor or absent response of parkinsonism to levodopa; early dysphagia and dysarthria; early onset of cognitive impairment including more than two of: apathy; impairment in abstract thought; decreased verbal fluency; utilization or imitation behaviour; or frontal release signs.

Exclusion criteria: A recent history of encephalitis; alien limb syndrome; cortical sensory deficits; focal frontal or temporo-parietal atrophy; hallucinations or delusions unrelated to dopaminergic therapy; cortical dementia of Alzheimer type; prominent, early cerebellar symptoms or unexplained dysautonomia; or evidence of other diseases that could explain the clinical features.

Once the classical syndrome is established, PSP should not be a difficult diagnosis to make. However in the early stages, the clinical picture may be incomplete and atypical phenotypes are now recognised complicating the diagnostic process. The true spectrum of the clinical phenotype is, however, broad (discussed in more detail in Chapter 3.2.1).

In the early stages of PSP, Parkinson's disease (PD) is a frequent misdiagnosis. When it becomes clear as in most cases that levodopa response is minimal, the differential often broadens to include multiple system atrophy, dementia with Lewy bodies or vascular Parkinsonism. It is in this early stage, when the more striking features of the disease such as the gaze palsy or the retrocollis may not be present, or when the disease presentation is atypical that making the correct diagnosis can be difficult. This

is reflected in the sensitivities of the different clinical diagnostic criteria which range from 13% to 83% (Litvan *et al.*, 2003).

1.2.4. Diagnostic difficulties in PSP

In addition to the early gait disturbance, postural instability and falls backwards, other early features include behavioural and cognitive changes, slurring of speech, visual symptoms and slowness. The disease onset is after the age of 40, but usually in the seventh and eighth decades of life (Lees, 1987; Steele *et al.*, 1964). In typical cases there is vertical supranuclear gaze palsy with the earliest neuro-ophthalmological finding being slowed vertical saccades (Stell and Bronstein, 1994). Supranuclear limitation of downgaze leading to falls when going downstairs and "the dirty tie syndrome" is characteristic.

There is a widening spectrum of clinical phenotypes documented in PSP and pathologically confirmed cases without an eye movement disorder have been described (Daniel *et al.*, 1995; Morris *et al.*, 2002; Revesz *et al.*, 1996). A number of other neurodegenerative and vascular pathologies have been associated with a clinical picture of PSP including diffuse cerebrovascular disease (Josephs *et al.*, 2002; Winikates and Jankovic, 1994), cerebral amyloid angiopathy with motor neuron disease (MND) (Weeks *et al.*, 2003), Cortico basal degeneration (CBD) (Litvan *et al.*, 1997), frontotemporal dementia with Parkinsonism linked to chromosome 17 (FTDP-17q) (Soliveri *et al.*, 2003), subcortical gliosis with prion protein deposition (Revesz *et al.*, 1995; Will *et al.*, 1988) and diffuse Lewy body disease (de Bruin *et al.*, 1992; Fearnley *et al.*, 1991; Nakashima *et al.*, 2003). Midbrain tumours (Silbert *et al.*, 1993), Whipple's disease (Averbuch-Heller *et al.*, 1999), neuro-syphilis (Murialdo *et al.*, 2000) and neuroleptic drugs (Campdelacreu *et al.*, 2004) are other rare causes of a similar clinical picture.

More recently, three cases have been reported, in which the combination of a bradykinetic-rigid syndrome and frontal cognitive impairment resulted in a diagnosis of PSP being considered during the disease course. In two of the three cases, PSP was the final clinical diagnosis. Neuropathological examination in these cases revealed typical histological appearances of fronto temporal lobar degeneration with ubiquitinated inclusions (FTLD-U) in one case, and fronto temporal lobar degeneration with motor neuron disease pathology (FTLD-MND) in two cases indicating that a clinical syndrome closely resembling PSP may be seen as part of the clinical presentation of these pathological disorders (Paviour *et al.*, 2004a).

The three cases described differed slightly from the clinical presentation of typical PSP. One presented with frontal behavioural change, which rapidly worsened within the first twelve months. Although formal neuropsychological examination was possible on only one occasion and was limited by dysarthria, this patient did demonstrate significant frontal executive dysfunction. Pathological examination in this case confirmed severe involvement of the frontal cortex. The neuropsychological findings in all three cases were, however, in keeping with a clinical diagnosis of PSP. The combination of severely slowed information processing and marked early executive dysfunction, are characteristic of PSP and have been suggested to help in differentiating it from some other degenerative dementias (Albert *et al.*, 1974; Litvan, 1994) (Chapter 3.5). In contrast, cortical dementia with dysphasia, meeting Neary criteria for semantic dementia is well recognised in FTLD-U (Rossor *et al.*, 2000) relating specifically to regional cortical atrophy. The neuropsychological profile may assist in the differential diagnosis, particularly if there is a suggestion of asymmetry but a predominant frontal dysexecutive syndrome can be found in any of the fronto temporal lobar degenerations, as well as in PSP and CBD. Despite this evidence for a

specific pattern of cognitive dysfunction, features of a "cortical dementia" such as dynamic aphasia have also been reported in PSP (Esmonde *et al.*, 1996).

The exclusion criteria for a clinical diagnosis of PSP include clinical features that clearly suggest an alternative diagnosis such as an alien limb and cortical sensory loss, favouring CBD, or early autonomic failure favouring a diagnosis of MSA. Hence while the current diagnostic criteria are useful, particularly in the context of recruitment to clinical trials, it is inevitable that even the most rigorous clinical diagnostic criteria will have limitations especially in the face of the increasing phenotypic variation of rare diseases such as FTLD-U and FTLD-MND.

1.2.5. Tau and genetics

PSP is a tauopathy. Tau is a microtubule-associated protein (MAP). It is found in neurones and in small amounts in glial cells. In the normal human brain, tau is distributed mainly in axons. By supporting cytoskeletal structure and sustaining axonal transport, tau plays a fundamental role in neuronal survival. In normal neurones, tau is soluble, binds to microtubules reversibly and has a rapid turnover. In neurodegenerative diseases such as PSP, tau loses its affinity for microtubules and collects as insoluble aggregates in the form of proteolytic resistant straight filaments.

PSP is characterised by the presence of neurofibrilliary tangles (NFT's) and neuropil threads (NT's) resulting from aggregation of hyperphosphorylated tau protein filaments. NFT's are basophilic and globose, their structure resembling a ball of string. However, flame shaped tangles (characteristic of Alzheimer's disease-AD) can also exist in PSP.

The human tau gene is located over 100kb on the long arm of Chromosome 17 and contains 16 exons. In the adult human brain, different isoforms of tau exist as a result

of alternate splicing of the tau mRNA transcript. Exons 2,3 and 10 are alternately spliced, allowing six different combinations of tau isoforms. The spliced products give rise to either three tau isoforms with three repeats (in the absence of the amino acid sequence coded for by exon 10) or three tau isoforms containing the amino acid sequence coded for by exon 10, with four repeats. Normal brain has three repeat tau isoforms. In PSP, hyperphosphorylated four repeat tau is abundant.

Recent studies have identified a genetic polymorphism in the tau gene associated with a greater than chance frequency with PSP. The most common allele in the general population is designated as A0. Other alleles are designated A1, A2 and A3. A significant over representation of the A0/A0 genotype in PSP has been described (Conrad *et al.*, 1997).

Two forms of the tau gene have been identified, the H1 and H2 genes. The H1 haplotype corresponds to the A0 polymorphism. A surprisingly large number of patients with PSP are H1/H1 homozygous, suggesting the A0/A0 phenotype may predispose to PSP. Two groups of patients with clinically typical and atypical, pathologically diagnosed progressive supranuclear palsy have been identified (Morris *et al.*, 2002). Of the clinically typical cases, 100% had the H1/H1 haplotype. Only 73.3% of the clinically atypical PSP group had the H1/H1 haplotype. The H1/H1 haplotype is not essential or sufficient to cause the disease but increases the risk of developing PSP. Exactly how the H1/H1 haplotype predisposes to PSP is unclear, but it may be related to splicing of exon 10.

The clinical genetics of PSP have been extensively studied (Rojo *et al.*, 1999). Among 133 first and second-degree relatives of cases of PSP in 12 families, 22 secondary cases of PSP were identified. No clinical differences were found between familial and sporadic PSP. The perceived pattern of inheritance in these cases was autosomal

dominant with variable penetrance. As described above, homozygosity for the tau A0 polymorphism has been described as a risk factor for developing PSP. It occurs in 55% of the population and the age-adjusted prevalence of PSP does not come close to this. In the study described above, a genomic search in two families revealed no linkage in the region of the tau gene (Chr17q 21-22). In addition, the A0 polymorphism had a similar distribution in affected and non-affected members of the families and so, the level of significance of the A0 polymorphism seems unclear. One large Spanish family has been reported in which four members with clinically typical PSP were screened and linkage to chromosome 1q31.1 identified. This region contains at least three genes, and their relevance in PSP is uncertain (Ros *et al.*, 2005).

1.2.6. Histopathology and macroscopic atrophy at postmortem

PSP is characterised by abnormal accumulation of tau protein in a number of brain areas but particularly the midbrain and basal ganglia.

The neuropathological characteristics of PSP (Hauw *et al.*, 1994) are legion and include macroscopic atrophy of the globus pallidus, subthalamic nuclei, brainstem and Brodmanns area 4. There may be dilatation of the IIIrd and IVth ventricle and aqueduct of Sylvius as well as de-pigmentation of the subthalamic nucleus (STN).

For the pathological diagnosis of PSP to be made accurately, the essential areas of the brain that must be studied are globus pallidus, putamen, caudate nucleus, STN, midbrain, cerebellar dentate nucleus and the pons and medulla. In addition, to exclude or diagnose other diseases, which may be mistaken for PSP, other areas such as the hippocampus and parahippocampal gyrus and motor, frontal and parietal cortices need to be examined.

PSP is characterised neuropathologically by numerous neurofibrilliary tangles (NFT's) and neuropil threads (NT's) in selected structures of the basal ganglia and brainstem together with variable neuronal loss and astrogliosis. These can best be identified by tau immunohistochemistry and silver staining.

In a typical case of PSP, NFT's are identified in the pallidum, subthalamic nuclei, corpus striatum and nucleus of Meynert as well as the brainstem (tegmentum, colliculi, peri-aqueductal grey, red nucleus, basis pontis, dorsal and median raphe nuclei and inferior olives) and dentate nucleus of the cerebellum. The prefrontal and precentral cerebral cortex may also be involved.

For a definitive final diagnosis of PSP, typical pathological appearances as well as a clinical history compatible with the diagnosis are required.

1.2.7. Correlation of pathology and clinical features

As already stated, NFT degeneration is greatest in the subcortical structures but PSP is also associated with varying degrees of cortical pathology. It has been suggested that the degree of cortical pathology has a bearing on the clinical phenotype and in a study of ten patients in whom the pre-central gyrus was the most severely affected cortical area, the density and location of NFT's appeared to correlate with the severity of the clinical symptoms possibly related to frontal cortical pathology (Verny *et al.*, 1996a). The presence of a supranuclear ophthalmoplegia may be correlated with brainstem pathology. However, the frontal eye fields and the parietal cortex are also known to be involved in the guidance of ocular movement and were found to be severely lesioned in these cases. The pathological involvement of the parietal cortex and frontal eye fields (FEF) was least severe in the only case without a supranuclear gaze palsy.

In another study linking regional pathological involvement and supranuclear gaze palsy in PSP, greater cell loss and NFT density in the nucleus raphe interpositus (NRI) was observed in patients with the typical clinical picture than in those without a gaze palsy suggesting this region of the pons is important in the generation of normal saccadic eye movements (Revesz *et al.*, 1996).

Braak et al (Braak *et al.*, 1992) found that in six patients with PSP, the three with significant cognitive impairment had greater severity of lesions in the entorhinal region of the cerebral cortex.

Kluin et al (Kluin *et al.*, 2001) concluded from a study of 14 patients with PSP that the hypokinetic aspects of dysarthria correlated well with the degree of pathology in the SN pars compacta and reticulata but not the striatum, STN or GP. The spastic and ataxic components of the dysarthria were not well correlated with pathological load in the cortex or cerebellum.

Hence the clinical features of PSP appear to be associated to some extent with the pathological involvement of definitive brain regions. If these clinical features are specific for PSP, then the neuronal loss in these areas ought to result in greater rates of atrophy in PSP than in other degenerative diseases and this may be apparent on MRI.

Consideration of clinicopathological studies is of great importance therefore when planning an imaging project involving a priori assumptions regarding which regions of the brain to study.

1.2.8. Clinical assessment of disease severity in PSP

A definitive clinical diagnosis of PSP is usually made, by a neurologist based on traditional clinical methods. Application of clinical diagnostic criteria are largely

restricted to clinical trials and research. Prognosis depends on relating the clinical stage of the disease and the presumed rate of disease progression.

Over time, experience of a particular disease enhances a physician's expertise in offering an informed opinion. This judgement is difficult to quantify and in clinical research, it is useful to incorporate clinical rating scales and bedside tests as quantifiable markers of disease severity. This aids comparison between individuals with the same disease, group comparison with other diseases and quantification of change over time.

1.2.8.1.Standardised clinical rating scales

There are a number of problems with the clinical scales currently in use. The Unified Parkinson's disease rating scale (UPDRS, see Appendix) is one of the most commonly used rating scales in Parkinson's disease (Fahn S, 1987). It consists of four subsets: Part I consists of 4 questions related to cognitive and psychiatric symptoms; Part II (often called the ADL component) comprises 13 questions related to common motor symptoms and their affect on disability; Part III is an objective motor examination; and Part IV asks questions related to complications of medications, such as motor fluctuations and dyskinesia. Responses to most questions are made on a scale from 0 to 4, with specific criteria for each value. It was designed to provide a measure of the signs and symptoms of PD in clinical practice and in research. As with all rating scales, its reliability and repeatability are paramount. As Parkinsonism is one of the clinical features of PSP and MSA, using the UPDRS motor subscale (Part III) to evaluate motor severity would seem reasonable. However rest tremor is not a common clinical feature of PSP, nor is limb rigidity and these are major components of the UPDRS motor subscale (UPDRS III). Axial rigidity and postural stability also

feature in the UPDRS III and a recent study suggests that it is a reliable and applicable scale for assessing most aspects of PSP function (Cubo *et al.*, 2000). Axial rigidity and gait abnormalities as well as limb bradykinesia were significantly related to the Hoehn and Yahr score (H+Y), see Appendix (Hoehn and Yahr, 1967) suggesting a good correlation between these factors. The H+Y score is possibly the most widely used and the simplest clinical rating score in Parkinsonism.

Other clinical features such as the supranuclear gaze palsy and cognitive dysfunction are not accounted for in the UPDRS III or the H+Y, although part I of the UPDRS allows for some subjective self-reporting of cognitive function. For these aspects of the disease, it is useful to have a bedside assessment as well as more formal approach to quantification of severity. Rating scales specifically designed for PSP are available but require further validation (Schrag *et al.*, 2005).

The mini mental state examination (MMSE) (Folstein *et al.*, 1975) is useful in any disease with a component of cognitive dysfunction, not least because it is so widely used and easy to administer, and most clinicians are very familiar with it. It is often used as a quick screening test for dementia and the score ranges from 0-30 with broad cut-offs of: >26 – no impairment; >20-26- mild dementia; <20- moderate dementia; and <12- severe dementia. The normal range is influenced by age and level of education and published normal scores are available but not often used. The MMSE score is heavily influenced by verbal and literacy skills. For neurodegenerative diseases with a frontal or sub-cortical pattern of cognitive dysfunction, the MMSE may be within the normal range despite significant impairment. With this in mind, the frontal assessment battery (FAB) has been developed and evaluated as a bedside test of frontal cognitive function (Dubois *et al.*, 2000) (see chapters 3.5 and 5.2.3). The FAB comprises 6 parts, testing various aspects of frontal cognitive function. It has a

maximum score of 18, with lower scores indicating progressive frontal cognitive decline.

The severity of gaze palsy can be quantified using electro-oculography, but this is time consuming, slightly invasive and not always routinely available. Alternatives utilising computer software and infra-red cameras are expensive. Bedside assessments of the severity of gaze palsy can be made using a simple rating scale that focuses on the saccadic eye movements. Judgements are made as to the speed (normal / mildly / moderately or severely slowed) and amplitude (normal / mildly / moderately or severely hypometric) of upward, downward and left-right saccades (see Methods chapter 4.2.1.5 where the scale is discussed in more detail). The scale was developed for use in the Neuroprotection and natural history in Parkinson plus syndromes (NNIPPS) study. There are no validated scales of gaze palsy published, and so this scale was applied to asses the severity of gaze palsy in the PSP patients. In this way, some standardisation of the severity and progression of the gaze palsy can be made, without recourse to expensive or invasive tests.

The cognitive impairment in PSP is an important component of the disease and while tests such as the FAB and MMSE provide some information regarding the level of dysfunction, more formal quantification may be more useful for detecting early abnormalities. The cognitive dysfunction arises as a consequence of basal ganglia pathology and damage to frontal-sub-cortical circuits. Many different neuropsychological tests have been developed to quantify the extent of cognitive dysfunction in PSP, MSA and PD. These are discussed in more detail in chapter 3.4.

For all of these tests, serial assessment may be used to determine change over time, but, more than in tests of motor function, "practice effects" need to be taken into account if there is a short interval between the tests.

All of these clinical scales for evaluation of disease severity are influenced by: floor and ceiling effects; variability in the test presentation and in some cases, the clinicians subjective assessment; and the patients performance on the day of the assessment which, may fluctuate according to a number of factors. This is regardless of the fact that on an individual level, the pathological disease process is relentlessly progressive. Hence these scales have limitations, particularly when used in isolation, when applied to quantify disease progression.

1.2.9. Conclusions

In summary, PSP is a neurodegenerative disease that at presentation can have phenotypic similarities with MSA-P and PD. This may lead to difficulties in discriminating the conditions. The established clinical diagnostic criteria (NINDS) have limitations and when applied late in the clinical course of the disease have a sensitivity of 85% (NINDS possible) and 34% (NINDS probable) compared to a sensitivity of 100% for an experienced clinicians diagnosis (Osaki *et al.*, 2004). Other studies of pathologically confirmed cases of PSP quote the sensitivity and specificity of an experienced clinicians final clinical diagnosis of PSP as 84% and 97% (Hughes *et al.*, 2002). However, earlier in the disease course identification of true positive cases of PSP either by an experienced clinician or by application of the diagnostic criteria is less accurate. The pathological disease process in PSP affects particular brain regions more than others. This regional distribution of pathology has a clinical correlate but current quantitative clinical rating scales may not be ideal for monitoring disease progression. In this regard, biomarkers are needed: neuroimaging, and MRI in particular may be useful as an adjunct to clinical assessment for increasing progression as well as supporting a clinical diagnosis.

2. Imaging the brain in PSP and bradykinetic-rigid syndromes

2.1 Introduction

Neurodegenerative diseases are characterised by progressive neuronal dysfunction and cell death. As neurons contribute to the volume of brain tissue, this neuronal attrition and subsequent cell loss, results in tissue atrophy. Macroscopic atrophy can be seen by the naked eye in the brainstem at post mortem in PSP and MSA, particularly in the midbrain in PSP. Modern neuroimaging techniques allow this macroscopic atrophy to be studied in detail and *in-vivo*.

2.2 Computed tomography in PSP.

The first study to demonstrate *in-vivo* midbrain atrophy in PSP, used pneumoencephalography (Bentson and Keesey, 1974). This investigation became obsolete with the advent of computed tomography (CT). Subsequently, numerous CT studies in small groups of patients with a clinical diagnosis of PSP were undertaken (Haldeman *et al.*, 1981; Masucci *et al.*, 1985; Masucci *et al.*, 1995). Characteristic changes identified in these cases included pontine atrophy, a dilated third ventricle, an enlarged quadrigeminal cistern and in some, a dilated aqueduct with a low-density area in the midbrain between the interpeduncular cistern and the aqueduct. Thin section CT studies of the brainstem discriminated some cases of PSP from PD on the basis of midbrain atrophy.

2.3 Magnetic resonance imaging.

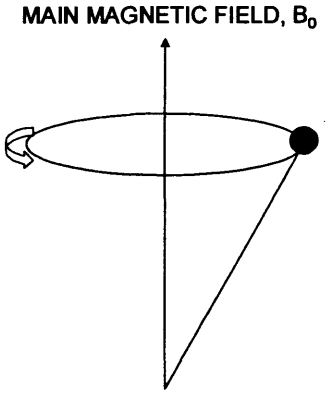
Magnetic resonance imaging (MRI) is established as a safe, non-invasive and high-resolution means of imaging the brain. MRI utilises the physical principle that protons in a magnetic field emit a radio signal following excitation by a radio frequency (RF) pulse.

In simple terms, the subject is placed within the magnetic field and protons in the water molecules of all body tissues, line up according to the magnetic moment of force. A RF pulse is then generated to excite the protons and when this is removed, the protons return to their previous alignment releasing energy, which is received as a signal and used to reconstruct an image of the body part being scanned. Routine MRI scanners generate magnetic fields (B_0) which vary from 0.3-3.0 Tesla (T) in terms of field strength. Unpaired protons become aligned with this field depending upon its strength and their thermal energy. Protons within a magnetic field "precess" (spin) in the direction of the field with a frequency (f_0) determined by the *Larmor* equation:

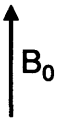
$$f_0 = gB_0$$

Where g is the constant defined by the magnetic property of the nuclei. When a RF pulse is transmitted at the resonance frequency of the proton, they absorb energy and move out of their original alignment to a higher energy level (decreasing longitudinal magnetisation) and such that they precess in phase with each other (establishing transverse magnetisation). When the RF pulse is turned off, the protons realign with the magnetic field and the energy they lose is emitted in the form of radio waves. The subsequent increase in longitudinal magnetisation gives rise to the T1 relaxation time; the loss of transverse magnetisation gives rise to the T2 relaxation time and is shorter than the T1 time constant (figure 2.1). The radio waves thus emitted can be converted into spatial information, as the magnetic field within the scanner is not uniform but is generated as a gradient, ensuring that emission of radio waves is additionally dependent upon the position of protons within the magnetic field.

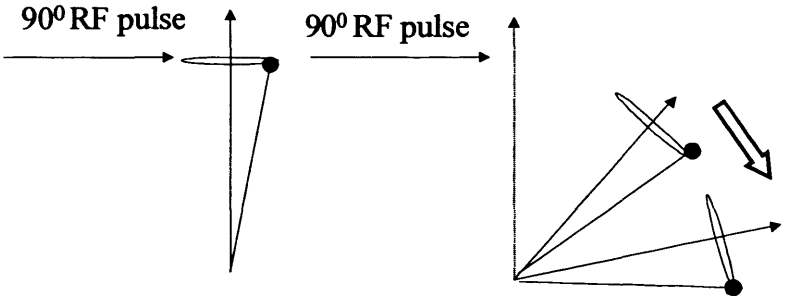
Figure 2.1 *Principles of MRI.*



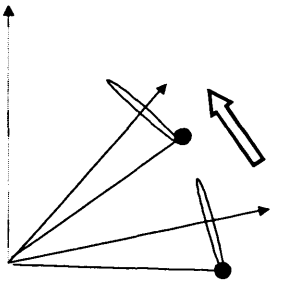
The nucleus of an atom (a single proton) spins or precesses within the magnetic field.



Increasing transverse magnetisation, decreasing longitudinal magnetisation



After the RF pulse is turned off, transverse magnetisation decreases (T1 relaxation) and longitudinal magnetisation increases (T2 relaxation)



Tissues may have different MR properties dependent on a number of factors including the number of free protons and how tightly bound they are. By altering acquisition

parameters, different T1 and T2 properties of tissues can be utilized to generate an image highlighting particular tissue types. The time to repetition (TR) is the time period between RF pulses and the longer this is, the more likely it is that protons will have a full recovery to their baseline state. A shorter TR allows the differential T1 properties of different tissues to be utilized in generating an image, as tissues may be in different states of excitation. It follows that short TR times improve T1 tissue contrast. Time to echo (TE) is the time between transmission of the RF pulse and collection of the signal. A short TE thus gives more T1 and less T2 contrast. In summary, a short repetition time leads to T1 weighting (dark CSF) and a long TE leads to T2 weighting (CSF appears white).

Alterations of TR and TE will have other effects on the scan and scan quality. Signal to noise ratio (SNR) will be increased with longer TRs, shorter TEs, increased voxel sizes and increased field strength. Scan time will be increased with longer TR, smaller voxel sizes and a larger field of view. Resolution is increased with smaller voxel sizes, with the consequence that this decreases SNR and increases scan time. Unlike CT, MRI data acquisition can be in different imaging planes and in three dimensions simultaneously (volumetric MRI). Volumetric MRI scans commonly consist of around one million 1-2mm^3 voxels (units of volume), each with a different tissue dependent contrast. T1 weighted volumetric scans have proven useful in quantifying atrophy.

2.4 MRI in degenerative bradykinetic rigid syndromes.

2.4.1. Introduction

None of the current diagnostic criteria for Parkinsonian syndromes (PD, MSA, PSP, CBD) include the results of any investigation as inclusion criteria, although certain investigation results may contribute to the exclusion of a diagnosis, for example evidence of early cardiovascular autonomic failure would exclude PSP as a clinical diagnosis.

In PD, which remains a clinical diagnosis based upon the presence of bradykinesia accompanied by at least one other characteristic clinical finding, (a resting tremor with a frequency of around 6Hz, the presence of rigidity and impaired postural reflexes) (Gibb and Lees, 1988), the signs and symptoms are usually asymmetrical at onset and when these typical findings together with a good response to dopaminergic treatment are present, it may not be necessary to image the brain. However, when the presentation is atypical, cross sectional T1 weighted MR imaging may be used in order to look for regional atrophy, more suggestive of one disease over another. T2 weighted MR imaging may also be used to exclude significant cerebrovascular disease or inflammation. MR imaging is also useful in excluding a secondary cause of Parkinsonism such as communicating hydrocephalus or a cerebral tumour.

An interest in the use of MRI in bradykinetic rigid disorders arose in the mid 1980's when it was recognized that MRI demonstrated brain iron deposition (Drayer *et al.*, 1986). While this is perhaps most striking in pantothenate kinase associated neurodegeneration (PKAN) due to PANK-2 mutations (Hayflick *et al.*, 2003), it is also a feature in degenerative diseases of the basal ganglia (Duguid *et al.*, 1986). Since the early 1980's other subjective and qualitative MRI findings have been proposed as helpful in discriminating Parkinsonian syndromes.

2.4.2. Qualitative MRI measurements as an aid to diagnosis

Brainstem atrophy and putaminal hypointensity, were found to discriminate atypical Parkinsonism (including PSP and MSA) from PD in early MRI studies (Stern *et al.*, 1989). However these MRI findings were unable to discriminate MSA from PSP reliably. A further study in the late 1980's supported this conclusion and added that MRI might have greater utility in discriminating MSA from PD as the MRI appearances in MSA were more uniform than in PSP (Savoiardo *et al.*, 1989). An extension of one of these studies reiterated that high field MRI is able to demonstrate iron in the brain and that Parkinsonian disorders can present an abnormal distribution of iron or other paramagnetic substances in the basal ganglia. In 20 patients with PSP, three experienced neuroradiologists evaluated the MRI findings noting atrophy of the midbrain in 11 cases, and that this atrophy was best seen on axial or mid-sagittal images. In 9 other patients, the midbrain size was considered to be borderline or normal. They also suggested that the third ventricle was disproportionately large compared to the lateral ventricles in PSP and that midbrain atrophy resulted in concavity of the posterior part of the floor of the third ventricle (Savoiardo *et al.*, 1994).

The ability of imaging criteria to discriminate PSP from CBD has also been studied. Midbrain atrophy and ventricular dilatation were confirmed as supportive of the diagnosis of PSP and asymmetric cortical atrophy supported the clinical diagnosis of CBD (Gimenez-Roldan *et al.*, 1994).

Other regions said to help discriminate PSP on the basis of qualitative assessments include regional atrophy of the corpus callosum on mid-sagittal MRI, which has been described as helping to distinguish PSP, AD and FTD with the pattern of callosal

atrophy (more anterior atrophy in PSP) consistent with areas of documented cortical hypometabolism (Yamauchi *et al.*, 1997).

In MSA, and particularly MSA of the cerebellar subtype (MSA-C), abnormalities on T1 weighted MRI were confined to the posterior fossa with predominant pontine and cerebellar atrophy.

In these studies, signal change in particular areas judged subjectively to be present or absent on T2 or proton density weighted imaging, was suggested as being supportive of the diagnosis of PSP or MSA. Reduced signal intensity in the substantia nigra with smudging of its margins towards the red nucleus was considered as supportive of a clinical diagnosis of PSP. A slight increase in signal intensity was noted in PSP on proton density images in the peri-aqueductal region.

In MSA, signal change in the basal ganglia was considered a more characteristic finding with low signal evident in the putamen and an occasional thin rim of hyperintensity noted at the lateral posterior putaminal rim, particularly in MSA-P. The so-called "*hot cross bun*" sign on T2 weighted MRI (figure 2.2) has been proposed as a useful clinical marker of MSA-C, (see p.49).

Figure 2.2 *T2 weighted MRI in MSA*
The "hot cross bun" appearance in the pons is clearly demonstrated on this T2 weighted MRI (arrowhead). A posterior to anterior line is clearly visible in the midline of the rostral pons and is bisected by a similar but less marked horizontal line. These signal changes are thought to reflect cell loss and gliosis.

This signal change in the pons is thought to be secondary to gliosis, which results in an anterior to posterior, and a left to right hyper-intense line. These specific changes were felt at the time to be pathognmonic of MSA-C, a statement which prompted other groups to publish examples of the MRI sign in other disease processes including CNS vasculitis, paraneoplastic syndromes, the spinocerebellar ataxias and myasthenia gravis (Burk *et al.*, 2001; Lo *et al.*, 2003; Muqit *et al.*, 2001; Murata *et al.*, 1998b).

Signal change on T2 or proton density weighted imaging may be expected in any region where the pathological disease process results in cell damage and cell death with reactive gliosis. The superior cerebellar peduncle is the main efferent from the cerebellum and is pathologically involved in PSP, with damage to the dentate nucleus

resulting in anterograde degeneration of the white matter tract as well as direct involvement of oligodendroglia in this region. The fibres of the SCP pass into the brainstem and decussate at the level of the inferior colliculi before synapsing in the red nucleus and the ventrolateral nucleus of the thalamus. A retrospective review of MRI scans in 9 patients with PSP (8 probable, 1 possible) compared with patients with PD and age matched controls subjectively identified signal change in 4 of 9 PSP patients and none in PD or healthy controls (Oka *et al.*, 2001).

Other studies of MRI in MSA compared to PD without incorporation of PSP patients have revealed various differentiating features, but largely support the findings detailed above (Bhattacharya *et al.*, 2002).

The main criticism of these studies is that most used crude and subjective methods for assessing the MRI abnormalities. Few employed any methods to control for inter-individual variation in brain size and images were often acquired on different MRI scanners, an important fact when interpreting intensity differences. From these studies, it seems that definitive and diagnostically supportive abnormalities were only seen in about 50% of patients with a clinical diagnosis of PSP. Some of these problems have since been addressed in imaging studies employing quantitative measures of atrophy.

2.4.3. Quantitative MRI measurements as an aid to diagnosis

Quantitative MRI measurements may have greater utility than qualitative assessments of MRI, in discriminating the degenerative bradykinetic rigid syndromes, particularly if statistical analysis demonstrates that a particular measurement discriminates PSP from the other diseases with a high degree of sensitivity and specificity and that the measurement is reliable and repeatable (see chapter 6.3.5 and 6.5.2).

Several studies have investigated the potential of MRI as an aid to diagnosis, setting out specific measurements of brainstem structures as diagnostic criteria (Asato *et al.*, 2000; Schrag *et al.*, 2000; Schulz *et al.*, 1999; Warmuth-Metz *et al.*, 2001).

One study demonstrated that patients with MSA-P and MSA-C as well as PSP could be discriminated from patients with PD and healthy controls on the basis of smaller brainstem volumes (Schulz *et al.*, 1999). The authors did not attempt to discriminate PSP on the basis of midbrain volume, or MSA on the basis of pontine volume and as such the conclusions have only limited clinical applications. A later study included qualitative as well as quantitative assessments (Schrag *et al.*, 2000) and concluded that the qualitative MRI features suggestive of a diagnosis of PSP over MSA-P were:

I. Signal increase and atrophy of the midbrain.
II. Thinning and smudging of the substantia nigra.
III. Atrophy and signal increase of the globus pallidus
IV. Atrophy of the red nucleus

In addition, supratentorial atrophy was seen in the temporal lobes and frontal lobes with infratentorial atrophy seen in the cerebellum, pons, and middle cerebellar peduncle. The conclusions stated that PSP could be differentiated from MSA-P using the following criteria:

I. Midbrain diameter <17mm (A-P diameter on axial scan).
II. Dilatation of IIIrd ventricle.
III. Signal increase in the midbrain.
IV. Frontal and temporal lobe atrophy.
V. Signal increase in the globus pallidus.

And also that MSA-P could be differentiated from PSP by the following:

I. Increased signal in the cerebellum and middle cerebellar peduncles.
II. Increased signal in the pons.
III. Dilatation of the IVth ventricle.
IV. Atrophy of the dentate nucleus.
V. Signal decrease in the globus pallidus.
VI. Hyper intense putamen or putaminal rim.

The only finding seen more frequently in CBD compared to MSA-P was global atrophy (however only a small number of CBD patients were included in the study).

The majority of these measures were qualitative and subjective, but the assessment of midbrain size in PSP was based on a linear measurement. A midbrain diameter of <17mm (on axial images) was very specific for PSP (96%) and had a positive predictive value of 80%. However its sensitivity was only 23%. A presumed correct diagnosis could still be made in the majority of cases of PSP and MSA-C.

Despite this, a significant number were not classified because either no characteristic changes were seen or because changes occurring in both PSP and MSA were seen, and only half the patients with MSA-P could be classified correctly.

The paper commented that 0.5T scans tended to misclassify patients with MSA as having PSP (19%) and that hence 0.5T scanners are unsuitable for this purpose. In total, 47% of patients could not be classified unequivocally.

Another group suggested similar quantitative MRI criteria for the differential diagnosis of PSP from MSA (Asato *et al.*, 2000). Again the quantitative measurement focussed on linear measurements of midbrain diameter, but this time on mid-sagittal images at the level of the posterior commisure. The authors concluded that in PSP the

rostral midbrain tegmentum (RMT) A-P diameter is smaller than 5.5mm (the lower 95% C.I for the normal range).

The pontine A-P diameter is often smaller than the lower limit of the normal range in MSA. This finding is supported by the results of another study, which concluded that a midbrain diameter of less than 14mm was found only in PSP (Warmuth-Metz *et al.*, 2001), however there was considerable overlap between midbrain measurements in the PSP and MSA patients.

Linear measurements may not be the best method of looking at regional atrophy, as they are so dependent upon the position of the scan slice and the orientation of the linear measurement. Asato *et al* ensured their imaging plane remained perpendicular to the axis of the brainstem in axial scans in order to partially overcome this problem. They also commented that compared to the methods employed in their study, volumetric or planimetric analyses would be less sensitive to atrophic changes in the midbrain as these measuring techniques include structures not affected in diseases such as PSP.

A limitation of the studies discussed so far is that they have made little attempt to account for the inter-individual variation in brain volume or head size that inevitably influences the interpretation of such quantitative analyses.

Intracranial volumes vary widely between individuals (95% C.I:1.2-1.7 litres) and systematically with height and gender (mean male TIV is ~12% greater than mean female TIV)

A ratio of linear measurements accounts in part for this variability and this has been employed in a recent quantitative study in which, PSP and MSA were reliably distinguished, suggesting future clinical usefulness (Oba *et al.*, 2005).

Volumetric MRI scans allow the true volume of a structure to be assessed, rather than an estimate based on a linear measurement. The study by Schulz *et al* is one of only a few volumetric studies to be published on PSP. Another relates to frontal atrophy and its correlation with behavioural change (Cordato *et al.*, 2002). PSP could be distinguished from PD and controls by whole brain atrophy, ventricular dilatation and frontal grey matter atrophy; however atrophy of the caudate nucleus seen in PSP patients in the study by Schulz *et al* was not replicated. There was no significant difference between PD and controls in any of the volume measurements.

This study concluded that frontal grey matter volume loss was the most useful for predicting diagnostic phenotype between PD and PSP. The authors concluded that, because no significant variability in frontal volume loss was seen despite significant variability in PSP disease duration, frontal atrophy occurs early in the disease. There were also associations between frontal volume on MRI and post mortem frontal volume as well as with the severity of the frontal behavioural changes during life.

This suggests that degeneration in other areas occurs or at least accelerates in the later stages of the disease. Serial studies at different stages of the disease would help to confirm or refute this.

2.4.4. Diffusion weighted imaging in bradykinetic rigid syndromes

Diffusion weighted imaging (DWI) is a rapid and practical imaging modality available on most clinical 1.5T MR scanners. The principles behind this imaging technique are discussed in more detail in chapter 8. DWI has been studied in MSA and PD (Schocke *et al.*, 2002) and this study found that that MSA can be differentiated from PD on the basis of higher regional apparent diffusion coefficients measured in the putamen, with a sensitivity and specificity of 100%. DWI has been

studied in a small group of patients with PSP (Ohshita *et al.*, 2000). As yet DWI has not been shown to be capable of differentiating PSP from MSA.

2.5 Correlation of imaging changes with pathology.

There have been few reports on the relationship between imaging *in-vivo* and pathological findings in the same patients at post mortem. If we presume that regional volume loss occurs as a result of pathological involvement of the brain tissue within that region, then in-vivo MRI findings should be associated with post mortem pathological involvement. Similarly, macroscopic atrophy at post mortem should correlate with the severity of pathological involvement microscopically.

Cordato (Cordato *et al.*, 2000) studied patients with PSP and PD as well as dementia with Lewy bodies (DLB) in order to quantify regional brain atrophy at post mortem and to correlate this with neuropathology. Frontal cortical volume and parietal cortical volume were specifically reduced in PSP. All three groups demonstrated significant medial temporal lobe involvement (amygdala) compared with controls. In contrast, only patients with DLB had hippocampal atrophy. In PSP the severity of atrophy was most marked in the globus pallidus internus, however structures such as the subthalamic nuclei and substantia nigra were not specifically studied.

The study largely represented late stage disease processes in pathologically confirmed cases. It would be important to determine the temporal progression of degenerative changes and their relationship to clinical signs in order to understand better their clinical significance. This cannot be done with pathological studies alone. However, an impression of the progression of atrophy or at least an idea of correlation between pathology and imaging could be obtained by undertaking MRI scans at two different time points during life.

Regional atrophy has been evaluated using T1 MRI of the brainstem on mid sagittal and axial images and compared to pathology in 8 cases of PSP subsequently confirmed at autopsy (Aiba *et al.*, 1997). The relationship between tau positive structures and atrophy was investigated, as was the histological nature of regions, which showed abnormal intensity on T2-weighted imaging. Midbrain atrophy with aqueduct dilatation on MRI was present in seven of the eight patients. Neuronal loss, a decrease in density of myelinated nerve fibres, gliosis and tau positive structures such as NFT's, glial fibrilliary tangles (GFT) and NT's were seen in these cases.

In patients who did not have atrophy on MRI scanning, histological cell loss was not significant. In areas of high signal intensity on T2 weighted imaging, histological analysis revealed severe neuronal loss with decreased density of myelinated nerve fibres, marked gliosis and abundant tau positive structures. Severity of atrophy was closely related to the density of the tau positive structures suggesting that atrophy on MRI reflects the degree of neuronal loss.

In a separate study, the most severely affected structures in the region of the midbrain were found to be the substantia nigra and the subthalamic nucleus (Yagishita and Oda, 1996).

In four PSP brains, the superior cerebellar peduncle was also noted to be small at post mortem and atrophy of this structure as it decussates may contribute to midbrain atrophy.

To conclude, there seems on the basis of these studies to be some degree of correlation between imaging findings in life and the distribution of pathology at post mortem in the brady-kinetic rigid syndromes. Large well controlled studies are lacking, however there is some support for the idea that regional signal change and

atrophy seen in life using MRI scanning represents relevant on-going pathological change in these conditions.

2.6. Serial imaging in bradykinetic rigid neurodegenerative diseases

Cross sectional imaging studies have limitations. There is large inter-individual variation in brain size and structure, the progression of atrophy can only ever be inferred from a single scan and it must be presumed that progression of atrophy is similar between individuals. Longitudinal studies use subjects as their own control and permit progression of atrophy to be quantified at an individual level. Hence longitudinal atrophy rates may distinguish early, mildly symptomatic neurodegenerative disease from "worried but well" individuals. Perhaps more importantly, rates of atrophy can be used as an outcome measure for disease modifying drug therapies which might reduce the rate of atrophy to that of normal healthy ageing. Finally, regions of atrophy may differentiate one neurodegenerative disease from another. This may be important if the diseases appear clinically similar but the underlying pathological mechanism is different, as future disease modifying therapies may work by combating protein dysfunction in one pathway but not another. For MRI to be applied as a means of supporting the diagnosis and monitoring disease progression: (1) it must be possible to compare serially acquired MRI scans from the same individual accurately and to apply robust techniques to measure volume change; (2) there must be an accurate understanding of the changes associated with normal ageing; (3) in addition to whole brain atrophy, some a priori assumptions about which regions to study need to be made.

In a cross sectional voxel based morphometric study of 465 *healthy* individuals (aged 18-79), age related grey matter loss was seen bilaterally in the insula, superior parietal

gyri and central and cingulate sulci, whereas the amygdala, hippocampi and entorhinal cortex exhibited little or no age effect (Good *et al.*, 2001). In a study using within subject registration of two MRI scans a year apart, of older disease free individuals (aged 59-85) age related grey matter loss in the hippocampus and in the orbitofrontal cortex was detected (Resnick *et al.*, 2000). These inconsistent results are likely to be due to differing image analysis methods.

Rates of global atrophy in healthy people increase gradually with age from an annual rate of 0.2% a year between the ages of 30-50 to 0.3-0.5% per year at ages 70-80 (Mueller *et al.*, 1998; Resnick *et al.*, 2003). This normal atrophy rate needs to be taken into account when comparing rates of atrophy in neurodegenerative disease states.

Very little longitudinal scanning has been undertaken in any of the Parkinsonian disorders. It has usually been limited to a few incidental scans in cross sectional studies and has not been carefully analysed or correlated with clinical or pathological information. Changes in MRI appearances over time in MSA have however recently been reported (Horimoto *et al.*, 2000; Horimoto *et al.*, 2002).

These studies looked at progressive atrophy in the different types of MSA compared to controls. The timing of the appearance of various qualitative features such as the "*hot-cross bun*" sign and putaminal slit sign in MSA-C and MSA-P respectively was noted. In the first study, atrophy was found to be greater in all MSA subtypes than controls and correlated more with duration of disease than age at onset. Atrophy rates were determined by two-dimensional area calculations on a single scan slice without registering the scans to standard space. The limitations of this method are discussed in more detail in chapter 6 and 9.

The second study concentrated on the appearances of certain imaging features at different stages of the disease. The hot cross bun sign in MSA-C was felt to represent atrophy and gliosis in the transverse pontine fibres and was present earlier in MSA-C than MSA-P. The putaminal hypointensity, representing iron deposition in that region was present earlier in MSA-P than MSA-C. No attempt was made in either study to correlate imaging features with the clinical phenotype. One useful clinical conclusion of the studies was that if putaminal hypointensity or a pontine "*hot cross bun*" sign are not present several years (~6) into the disease process, then a diagnosis of MSA can be considered unlikely.

In another longitudinal study of clinical features in 230 patients with MSA, 139 subjects had MRI scans reviewed qualitatively. Whilst predominantly cross sectional, this study demonstrated MRI findings in a large number of patients at different stages of the disease process, giving insight into the evolution of imaging features over time (Watanabe *et al.*, 2002).

2.7 Image registration

Cell loss that accompanies histological neuro-degeneration can be examined in vivo as atrophy on MRI. Because of the wide normal variation in cerebral structures, distinction of atrophy associated with neurodegenerative conditions from changes associated with normal ageing is difficult using imaging at a single time point, especially in the early stages of a disease process.

Measurement of atrophy on a single scan does not make full use of the structural detail available as small changes can be masked by the biological variability in normal individuals. In serial imaging of individuals, the first scan may be used as a

reference point for future change, hence allowing intra-individual change over time to be assessed.

2.7.1. Technical considerations of serial imaging

The accurate assessment of inter-individual differences or within-subject change over time is aided by matching scans positionally so that they are in a common spatial framework. This needs to be performed post acquisition, as with serial scanning, subjects cannot be placed in as similar a position within the scanner to allow meaningful comparison without post-acquisition registration. Even with very good post acquisition registration, it is still desirable to try and position patients as consistently as possible. The reason for this is that there are both spatial and intensity inhomogeneities within the scanners field of view. Modern short bore scanners are particularly prone to non-linear geometric distortions. As a result, scanner manufacturers have put considerable effort into correcting these distortions. Positioning may be a particular problem in elderly patients with limited mobility or in diseases which affect posture such as Parkinson's disease, PSP and MSA.

There are also problems due to inevitable changes in the magnetic field that occur over time due to scanner gradient changes. These changes can result in fluctuation in voxel size, complicating comparison of baseline and repeat scans (Whitwell *et al.*, 2001).

Every effort must still be made to position reproducibly the subject in the scanner, and to minimise hardware and software changes during longitudinal studies. With consistent image acquisition, post acquisition registration can permit precise assessments of change.

The goal of post acquisition registration is to align regions of interest on a baseline scan to the same region in follow up studies, so that the voxels representing the same underlying anatomical structures are superimposed. This allows for direct comparisons of either global or regional brain structures. For the purposes of comparing structural scans and calculating rates of atrophy, rigid body registration (in which the baseline and repeat scans are considered to be rigid structures that must be matched onto one another without structural change) is the most commonly used technique.

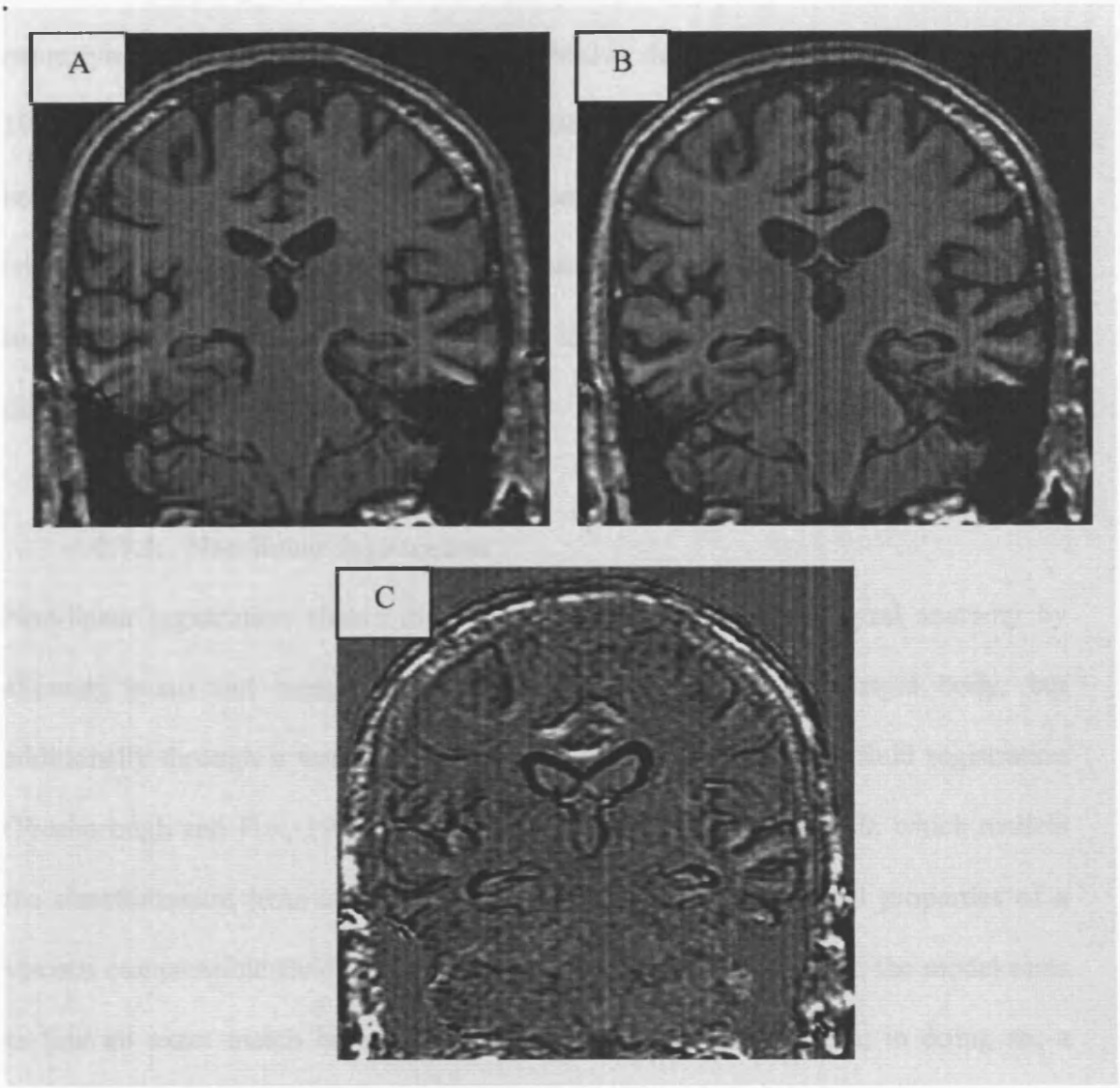
2.7.2. Rigid body registration

A number of techniques have been designed to perform rigid body registration (for a review see (Ashburner *et al.*, 2003). The most basic method is to identify a number of corresponding anatomical points on two consecutive scans and to determine the transformations that align these points accurately on top of one another. A minimum of three linearly independent points are required to determine the unique three dimensional co-ordinate system allowing superimposition. The more points that are determined on the scan, the greater the registration accuracy. However, manual definition of such points is time consuming and accuracy is limited by the ability of the user to define reliably the anatomical structures. Furthermore, the technique demands the existence of several invariate points and such lack of variation cannot be assumed in brains undergoing neurodegeneration. Several automated techniques have been designed to allow accurate scan matching (Freeborough *et al.*, 1996; Woods *et al.*, 1998). These algorithms compare signal intensities from all voxels within the brain. By matching signal intensities from the voxels of the repeat to the baseline scan, it is possible to improve significantly registration of intra-subject MR scans. The aim is to superimpose voxels of similar signal intensity, thereby using all the

information within a scan to perform the matching process. This voxel intensity based method of registration (Freeborough *et al.,* 1996) involves: (1) segmenting the scan to remove non brain tissues; (2) determining and applying a series of rotations and translations to match a repeat scan to the baseline. Translations and rotations are applied in each direction within and around the x, y and z axes producing a so-called 6 degrees of freedom (*6dof*) registration. This process can be extended to allow linear scaling to a small extent in each of the three planes (nine degrees of freedom registration -*9dof*). This theoretically allows for correction of inconsistencies in voxel size that may occur between the two scans. Once registered, a difference image can be created, so change between scans can be visualised (figure 2.3).

This technique or its variants has been extensively validated (Freeborough *et al.,* 1996; Gunter *et al.*, 2003; Paling *et al.*, 2004) and is now widely used to compare serially acquired scans.

Figure 2.3 *Registered coronal T1-weighted MRI*



A-Baseline image; B-Repeat scan of the same subject approx 1 year later, registered to baseline. Note the good anatomical correspondence despite the atrophy that has occurred over the interval.; C-Difference image (scan A subtracted from scan B), revealing tissue loss (dark areas) around the ventricles.

2.7.2.1.Atrophy calculations

Once two serial scans have been registered to one another, outlining of the brain or brain structures can then be performed on both baseline and repeat scans. When performed on sequential slices of a volumetric image, volumes of certain brain structures can be accurately assessed and atrophy rates calculated. Outlining small structures by hand is inevitably associated with measurement error however (Chapter 6).

The brain boundary shift integral (BBSI) has been validated as a robust tool for quantifying global brain atrophy in degenerative dementia (Fox and Freeborough, 1997; Freeborough and Fox, 1997). The BBSI calculates the shift at the brain/CSF boundary at every point across the three dimensional registered scan pair, the sum of which approximates closely to the brain volume lost; being a direct measure this leads to a reduction in the error associated with indirectly comparing two scans. This is discussed in more detail in chapter 9.

2.7.3. Non-linear registration

Non-linear registration allows for a more accurate matching of gyral anatomy by allowing scans not merely to match through movement as a rigid body, but additionally through a warping proceedure. One such approach is fluid registration (Freeborough and Fox, 1998) discussed in more detail in chapter 10, which models the transformation from one scan to another based on the physical properties of a viscous compressible fluid. Following a *9dof* rigid body registration, the model aims to find an exact match between the baseline and the second scan; in doing so, a displacement vector is generated for each voxel within the image, with the degree of stretch or contraction demonstrating expansion or atrophy over the period; this can be demonstrated as an overlay image (voxel compression map) providing a means of demonstrating the pattern of regional atrophy throughout the brain, occurring between two scans without the need for a priori decisions about which structures should be assessed.

This allows an individual's pattern of progressive atrophy to be visualised. An extension of this is to critically compare at a group level, fluid-registered scans from individuals within a statistical framework (statistical parametric mapping-SPM,

(Friston, 1995), allowing consistent patterns of atrophy within *groups* of patients to be demonstrated (see Chapter 11). SPM analysis of fluid registered images incorporates the neuroanatomical sensitivity of within-individual registration and the power of group comparisons done with between-individual registrations. Fluid registration has been shown to be useful in propagating outlined hippocampal volumes forward from a baseline scan to a repeat scan, thus potentially increasing accuracy and saving operator time in making hippocampal measurements (Crum *et al.*, 2001). Other computational methods for determining atrophy patterns in an unbiased fashion have also been developed and a comprehensive overview of these techniques has been published (Ashburner *et al.*, 2003).

Whatever method is used it needs to have high spatial resolution and clear differentiation between tissue types. Normally, three dimensional, high resolution, T1 weighted MRI acquired with conventional 1.5T MR scanners and $1mm^3$ voxels across the cranium provide sufficient detail and contrast.

As MR scanners regularly undergo substantial hardware and software modifications and as these can substantially alter image quality and intensity contrast, the same scanner and MR acquisition parameters should be used for baseline and repeat scans. As with any other method of measuring changes in brain volume over time, these methods are also sensitive to subtle brain changes that are not directly disease related, such as hydration and confounded by movement artefact.

2.8. Brain atrophy as a marker of disease progression

The last five years have seen the introduction of new drugs aimed at modifying disease progression or alleviating the symptoms of neurodegenerative disease, for

example riluzole in motor neurone disease (MND) and donepezil, rivastigmine galantamine and memantine in AD.

Drug trials require accurate clinical diagnosis to ensure purity or at least some homogeneity of the study cohort. In the NNIPPS (Neuroprotection and Natural History in Parkinson-Plus syndromes), a randomised controlled trial looking at the effect of Riluzole in PSP and MSA, diagnosis depends on consensus clinical criteria and there is a lack of well validated methods, which are not potentially influenced by symptomatic drug effects, with which to monitor disease progression. Modelling of the trial can in part overcome this problem but generally, clinical evaluation of progression suffers from being difficult to distinguish from symptomatic effects without complicated trial design, or else is not easily quantifiable (e.g. the severity of gaze palsy). In cases with cognitive decline, psychometric data are potentially useful diagnostically and can be used to measure treatment effects, but require large numbers of cases which is not always practicable for these conditions.

There are at present no biomarkers to diagnose PSP and importantly for clinical trials, no reliable quantitative measure of disease progression.

An objective, quantitative measure of disease progression is badly needed for evaluation of novel drug therapies. This prospective study will identify volumetric changes over time, which may prove useful as a quantifiable pragmatic measure of disease progression for future clinical trials.

2.8.1. Potential problems with current outcome measures

The utility of many neuropsychological tests and rating scales of clinical symptoms, has been demonstrated in clinical studies in PSP and in the other bradykinetic rigid syndromes (Cubo *et al.,* 2000). The Unified Parkinson's disease rating scale (UPDRS) has been shown to correlate with the degree of ^{123}I-beta-CIT binding to striatal

dopamine uptake sites in PD (Brucke *et al.*, 1997) suggesting that the score might correlate with nigral degeneration. However in disease modifying drug trials, the power of these tests to quantify disease progression rather than simply the symptomatic effect of drug intervention is unclear and striking disparities between the rating scale and functional imaging changes have occurred.

In a *randomised withdrawal design,* patients are initially randomised to either active drug or placebo. The trial continues for a significant duration for the drug to demonstrate efficacy over placebo in terms of the marker of disease severity (such as a clinical rating scale). Subsequently, those randomised to active drug in the initial phase are further randomised to either continue on the active drug or to receive placebo. If at the end of this second phase, those receiving placebo (second phase only) maintain their difference from those receiving placebo through both segments, a disease modifying effect is inferred if the second phase has been long enough. If in the second segment, those receiving placebo deteriorate in their clinical scores to the level of those who have received placebo throughout, then a purely symptomatic effect is suggested.

In a *randomised start design*, patients are initially randomised to either active drug or placebo. At the end of the first period, those initially receiving placebo are randomised to receive active drug or placebo. If at the end of this second segment, the group initially receiving placebo but then switched to the active drug, catch up those who have recieved the active drug from the start, a symptomatic effect is inferred. If the difference between the groups is maintained, then a disease modifying effect is postulated.

These studies, by their nature, need to include large numbers of patients and continue for at least a year.

In an effort to reduce the completion time and patient numbers, surrogate markers of disease progression are needed urgently.

2.8.2. Surrogate markers

"A surrogate endpoint for a clinical trial is a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives. Changes induced by therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint." (Temple, 1999). The properties of an ideal surrogate endpoint should be:

1. It must not merely be a correlate of the true clinical outcome
2. The effect of an intervention on a valid surrogate endpoint must reliably predict the effect of the clinical outcome of interest
3. The treatment effect on the clinical outcome should be explained by its effect on the surrogate marker (Fleming and DeMets, 1996).

2.8.2.1. Successful surrogate markers

An example of a surrogate marker in a clinical disease setting is the CD4 count in HIV infection. The HIV pathogen attacks CD4 cells. Reduction in CD4 cells leads to immuno-suppression, which leads to morbidity and mortality. Drugs to treat HIV should increase the CD4 count, which should reduce the morbidity and mortality rate. Hence the CD4 count is used as a surrogate marker for mortality (the true outcome measure) in HIV treatment trials.

2.8.2.2. Failed surrogate markers

There are a number of reasons why surrogates may fail. Again these have been highlighted by Fleming and DeMets (Fleming and DeMets, 1996):

1. While a surrogate may be a marker of the disease process, its progression may involve a different pathological process. Hence intervention that affects a surrogate, may not affect the true clinical outcome.

2. A disease may progress through several causal pathways, not all of which are mediated through the surrogate. Thus an intervention affecting the surrogate may not fully predict the effect on the true clinical outcome.

3. An intervention might affect a pathway leading to the true clinical outcome without impacting on the surrogate.

4. The intervention might, through unintended mechanisms, affect the true clinical outcome independent of the disease process.

2.8.3. Serial imaging as a marker of disease progression

Using brain imaging as a surrogate marker of disease progression in neurodegenerative diseases is a concept first proposed in 1996 (Smith and Jobst, 1996).

No disease modifying therapy is available in PSP. A claim for dopamine agonists as a neuroprotective agent in PD has been mooted on the basis of positron emission tomography (PET) and clinical outcome measures (Whone *et al.*, 2003). This data, however cannot exclude the possibility that symptomatic benefit entirely explains the results. In the absence of any proven disease-modifying drug, it is not possible to determine the effect of a drug on disease progression using MRI. However, there is increasing interest in Alzheimer's disease (AD), to support the use of MRI measures of progression, as a viable surrogate marker.

2.8.3.1. Biological plausibility

In PSP and MSA, cross sectional imaging studies imply atrophy in specific brain regions compared to normal controls. As very few serial imaging studies have been

undertaken in these disorders, it remains uncertain whether clinical disease progression is reliably associated with an increased rate of brain tissue loss.

However in AD, numerous studies have confirmed that excess cerebral atrophy occurs prior to the onset of symptoms in sporadic AD (Rusinek *et al.*, 2003) and appears an inevitable consequence of disease progression (Fox and Schott, 2004). Studies in AD have also demonstrated correlations between the rate of cerebral atrophy in life using MRI and underlying pathological change at post mortem (Silbert *et al.*, 2003). Cases of serial imaging correlations with pathological findings at post mortem in PSP and MSA are described in chapter 3.5.2.

It is reasonable to hypothesise that in PSP, as in AD, excess rates of atrophy are: (1) a consequence of neurodegenerative disease; (2) that the rate of clinical disease progression depends at least partly on the rate of brain tissue loss (atrophy); (3) that the clinical phenotype depends on the distribution and tempo of that tissue loss (regional atrophy) and; (4) the rates and regions of atrophy can be reliably quantified using serial MRI.

2.8.3.2. Feasibility

In AD, the rates of atrophy measured using serial MRI have been calculated, and the number of patients required to power a drug trial can predicted from this. The sample sizes required to power a similar study in PSP are unknown as serial imaging studies have not been undertaken. It would be reasonable to assume, however, that the rate of total tissue loss is likely to be lower in PSP than in AD and for this reason, the number of patients required to power a drug trial would be greater. Whether measuring regional rather than whole brain atrophy in PSP would provide a method of monitoring disease progression with a smaller sample size is unknown. In AD, it has

been demonstrated that similar numbers are required regardless of whether whole brain or regional atrophy rates are used (Jack, Jr. *et al.*, 2004).

2.9 Conclusions

All neurodegenerative diseases require brain histopathological examination for definitive diagnosis. Whilst the classical clinical features may differentiate neurodegenerative diseases, heterogeneity and overlap syndromes present considerable difficulties. Furthermore, the severity of clinical features may not always accurately predict the stage of the disease or its rate of progression. A good example would be the supranuclear gaze palsy which may occur early and be severe or may be completely absent in PSP.

The cell loss that accompanies histological change in these neurodegenerative diseases can be visualised using MR imaging. Regional atrophy results in features on cross sectional MRI that have been proposed as helping to discriminate PSP from MSA, PD and healthy controls as well as from other degenerative diseases. However, many of these regions have not been quantitatively examined. Based on knowledge of the distribution of pathological changes at post mortem, there are also other regions which may prove to be better candidates for discriminating PSP. Imaging studies that make no a priori assumptions about where differences between diseases occur may help to identify new regions to study in more detail.

Measures of atrophy over relatively short periods of time (a year for example) are more likely to be linear. As progressive excess atrophy rates appear to be an inevitable feature of the neurodegenerative disease process closely allied to the underlying pathology of the disease, that progressive atrophy is a biologically plausible marker for disease progression. Hence any trial suggesting that an

intervention, compared to placebo produced a significant reduction in brain atrophy rates would imply a disease modifying effect in addition to any cognitive, behavioural or motor effects demonstrated using clinical rating scales of disease severity.

3. Designing a prospective serial imaging study in bradykinetic-rigid syndromes: Preliminary studies and methods overview.

3.1. Hypothesis and aims of the study

Whole brain atrophy rates in PSP and MSA might be similar given that these neurodegenerative diseases have a similar age at onset; disease duration and both have a relentless progression over 6-9 years to death. Therefore in addition to comparing rates of whole brain atrophy in PSP, MSA-P, PD and healthy controls, the aims of this study include comparing regional rates of atrophy in PSP, MSA-P, PD and healthy controls.

The differential diagnosis of PSP, MSA-P and PD can be difficult as the clinical picture may in some cases be similar, particularly in the early stages of disease. MRI may help improve the accuracy of diagnosis during life and provide correlations with clinical features, clinical rating scales and neuropsychological scores. This study will clarify the association of clinical disease progression with atrophy and look for new MRI based diagnostic markers in PSP. It will also use novel image analysis techniques to determine the patterns of progressive global brain atrophy occurring in PSP, MSA-P and PD without making a priori assumptions about where that atrophy might occur.

In this way, this study aimed to determine how MRI, as a method of determining brain atrophy and disease progression, and as a diagnostic aid can best be applied to PSP.

3.2. Planning the clinical assessments

In the original paper describing progressive supranuclear palsy, the authors predicted that "further observations may (in the future) broaden the clinical spectrum of the

disease" (Steele *et al.,* 1964). Subsequently cases presenting with pure akinesia without rigidity (Matsuo *et al.*, 1991; Verny *et al.,* 1996a), gait freezing (Matsuo *et al.,* 1991), rest tremor (Birdi *et al.*, 2002; Verny *et al.,* 1996a), isolated dementia (Davis *et al.*, 1985; Masliah *et al.*, 1991), parkinsonism without dementia (Birdi *et al.,* 2002; Davis *et al.,* 1985; Verny *et al.*, 1996b), limb apraxia and asymmetric parkinsonism (Motoi *et al.*, 2004) as well as a number of cases dying without recorded evidence of the distinctive supranuclear ophthalmoplegia (Daniel *et al.,* 1995; Davis *et al.,* 1985; Dubas *et al.*, 1983) have been reported. Despite these observations the current operational criteria for the diagnosis of progressive supranuclear palsy (PSP) includes very few clinical features that were not recognised in the original description (Litvan *et al.,* 2003).

Attempts have been made to embrace the broader clinical features that are now well documented in pathologically confirmed PSP, but are not part of the operational clinical diagnostic criteria. An imprecise classification has evolved which includes typical and atypical clinical subgroups. The definitions of these subgroups have varied and have usually been applied retrospectively. They include: the presence or absence of supranuclear gaze palsy (Birdi *et al.,* 2002; Daniel *et al.,* 1995); the presence or absence of a diagnosis of PSP in life (Gibb *et al.*, 2004; Morris *et al.,* 2002); the application of retrospective diagnostic criteria (Morris *et al.,* 2002); and the presence or absence of early bulbar signs or falls (Nath *et al.*, 2003). Several authors have found genetic, prognostic or pathological differences between these putative clinical groups (Daniel *et al.,* 1995; Morris *et al.,* 2002; Nath *et al.,* 2003).

That significant prognostic factors in a progressively debilitating neurodegenerative condition have been found is not surprising. However, the clinical features that have been reported to purport the greatest prognostic significance are also the classic

clinical hallmarks described in Richardson's original reports (Richardson *et al.*, 1963; Steele *et al.*, 1964) and are embraced by the accepted operational diagnostic criteria (Litvan *et al.*, 2003). Accordingly the absence of supranuclear gaze palsy, early falls and early bulbar dysfunction, with a positive response to levodopa conveys a good prognosis, but patients presenting in this way are much less likely to be diagnosed with PSP (Daniel *et al.*, 1995).

There are no non-invasive biological markers for the antemortem diagnosis of PSP (Litvan *et al.*, 2003), and the "definite" diagnostic category has traditionally been reserved for cases that are pathologically confirmed (Litvan *et al.*, 1996a; Litvan *et al.*, 1996b). The cases that remain undiagnosed in life make up at least 20% of the pathologically diagnosed cases of PSP (Hughes *et al.*, 2002) and most have unusual or atypical clinical pictures (Daniel *et al.*, 1995; Litvan *et al.*, 1999; Morris *et al.*, 2002). Hence in order to study the correlation of clinical and imaging features properly during life, it is important to have a clear understanding of the heterogeneity of PSP. With this in mind, all cases with a pathological diagnosis of PSP in the Queen Square Brain Bank for Neurological Disorders, were reviewed and the clinical features documented in a standardised manner.

3.2.1. Clinical features of PSP cases in a regional brain bank

3.2.1.1. Patients

One hundred and three consecutive cases of pathologically diagnosed PSP, donated from all over the UK between 1988 and 2002 and stored in the Queen Square Brain Bank (QSBB) at the Institute of Neurology, UCL were reviewed. The diagnosis was made according to the NINDS-SPSP criteria (Litvan *et al.*, 1996b), and was retrospectively applied to those cases acquired prior to the publication of the criteria. All patients had been assessed during life by a hospital specialist (neurologist or

geriatrician). Some of these cases have been included in previous clinical and pathological reviews from the QSBB (Daniel *et al.*, 1995; Gibb *et al.*, 2004; Hughes *et al.*, 2002; Morris *et al.*, 2002; Osaki *et al.*, 2004).

3.2.1.2. Clinical data collection

A systematic chart review was performed on all the case notes. Specifically the comprehensive case notes of the family doctor and all of the correspondence between the family doctor and the medical specialist were reviewed. When available, medical inpatient notes, inpatient consultations and emergency room admission notes were also scrutinised. A clinical data sheet was designed to record the presence or absence of clinical features either early in the disease course (within two years of first symptom onset) or at anytime during the disease. Symptoms were recorded as being *absent* if not reported, and clinical signs were recorded separately as *unknown* if they were not specifically mentioned in the notes. Where conflicting clinical features were recorded the findings of the neurologist were used.

Definitions were: Age at onset: Age, in years, at the time of the first reported symptom considered to be attributable to PSP or Parkinsonism. Duration: Time between the age at onset and the age at death. Falls: The presence of any report of falls. Bradykinesia: The presence of any mention of bradykinesia or motor slowing. Cognitive decline: The presence of any perceived cognitive decline, either by the patient, patient's relative or the treating doctor. This included annotation of difficulty in concentration, reports of intellectual functional decline and comments by the family of mental slowing but did not include affective disorders. No patients underwent formal neuropsychological testing in the first twelve months of their disease. Speech disturbance: The recording of any alteration in speech quality compared to speech prior to disease onset. Dysphagia: A record of any swallowing abnormality.

Asymmetric Onset: If there was a clear difference between the signs on the left and the right asymmetry was recorded as being present. This included asymmetry of tremor, rigidity, bradykinesia or functional decline. It did not include specific tasks such as writing and using tools. Tremor: The recording of any tremor. Rigidity: The recording of axial or peripheral muscle rigidity, extra-pyramidal and pyramidal rigidity was not differentiated. Impaired Postural Reflexes: The presence of this sign was only recorded if specifically mentioned in the clinical notes. Supranuclear Gaze Palsy: The specific recording of restricted range of eye movement in the vertical plane. Impaired Saccadic or Pursuit Movements: The specific recording of abnormal saccadic or smooth pursuit eye movements. Other visual symptoms: The recording of other visual symptoms not explained by the presence of gaze palsy or impaired saccadic or pursuit movements, which evolved during the disease course. Symptoms include painful eyes, dry eyes, visual blurring, diplopia, blepharospasm and apraxia of eyelid opening. Extra axial-dystonia: The presence of dystonia in any body part apart from trunk and neck. Pyramidal signs: Pathologically brisk reflexes and/or extensor plantar response(s). Autonomic dysfunction: Either abnormal autonomic function testing or documentation of any two of: urinary urgency, frequency and nocturia without hesitancy; chronic constipation; postural hypotension; sweating abnormalities; erectile dysfunction. Dyskinesia: The presence of chorea associated with levodopa therapy. Response to levodopa: The patient and clinician's interpretation of improvement was assessed from the case notes and in some cases from the completed Parkinson's Disease Society Brain Bank Annual Assessment (PDSBB) Forms. A self reported improvement of more than 30% coincident with the introduction of levodopa was recorded as being a positive response. This degree of response was graded by a 4 point scale modified from the PDSBB annual assessment

forms: 1 - nil, or slight response (<30% improvement), 2 - moderate response (30-50% improvement), 3 - good response (51-70% improvement) and 4 - excellent response (71-100% improvement).

3.2.1.3. Statistical methods

A complete data set of clinical variables each coded as present or absent was available from 29 cases only and these were entered into a principal components analysis in order to summarise the information. Other cases had missing data on one or more variables and could not be included in this analysis. Statistical analysis was performed using SPSS for Windows (version 12.0.1). The principal components analysis was performed using clinical data from the first two years of disease. The first two principal components were selected and "Varimax" rotation performed (this rotation maximises the number of variables that have high loadings on each factor). The loadings of each variable on both of these components were plotted against each other. The plot was examined and two groups of variables in different areas of the plot were selected. A between groups, hierarchical cluster analysis using squared Euclidean distance measures was performed to check that these sets of clinical features grouped together. Those cases that exhibited a greater number of the characteristics from set 1 than from set 2 were deemed to fall within group 1, and vice versa for group 2. A cross tabulation was used to examine how many cases had a similar number of characteristics from the two variable sets. This was then applied to the cases that were excluded from the principal components analysis. Where there were missing data, the number of positive results found were scaled up (e.g. multiplied by 5/3 where status on 3 out of 5 characteristics was known). Clinical characteristics of the patients were compared between the two groups, and

Table 3.1 *Clinical features in 103 pathologically confirmed cases of PSP*

	First 2 years				Throughout disease			
Clinical Feature	Present early	Absent early	Not recorded	Present in total cohort %	Present ever	Absent ever	Never recorded	Present in total cohort %
Falls	61	40	2	60	93	8	2	94
Bradykinesia	74	25	4	75	99	0	4	98
Tremor	20	79	4	20	23	76	4	23
Cognitive Decline	28	68	7	29	72	25	6	74
Non-specific visual symptoms	20	77	6	21	48	49	6	50
Rigidity	43	53	7	42	95	2	6	98
Dysphagia	6	89	8	6	64	29	10	69
Speech disturbance	38	60	5	39	83	12	8	87
Impaired postural reflexes	49	30	24	62	84	2	17	98
Supranuclear gaze palsy	19	31	53	38	75	7	21	91
Abnormal pursuit or saccades	21	27	55	44	70	6	27	92
Extra-axial dystonia	7	89	7	7	25	70	8	26
Pyramidal signs	8	88	7	8	18	79	6	19
Cerebellar signs	0	10	3	0	1	98	3	1
Autonomic dysfunction	0	100	3	0	2	97	3	2
Cortical sensory loss	0	97	4	0	1	97	4	1
Asymmetric Onset	26	67	10	28				
Levodopa response					29	62	12	32
Levodopa induced dyskinesia					4	92	7	4

PSP = progressive supranuclear palsy

significance was calculated using Student's t-test for normally distributed data, and chi-square or Fisher's exact test for binary data.

3.2.1.4. Results

3.2.1.4.1. Clinical features

Clinical data was available for 103 pathologically confirmed cases of PSP, of which 65 were male (63%). The clinical diagnosis of PSP was made during life in 71 cases (69%), 24 (23%) were diagnosed with idiopathic Parkinson's disease (PD), 2 (2%) with atypical PD, 2 (2%) with multiple system atrophy, 1 (1%) with corticobasal degeneration and in 3 a final clinical diagnosis was not established. Neurologists made the final clinical diagnosis in 87% of cases and correctly diagnosed PSP in 72%.

The mean (SD) age of onset was 66.4 (12) years, age at death 73.5 (7.5) years, and

disease duration was 7.0 (3.7) years. The clinical features of these patients are summarised in Table 3.1.

An asymmetric onset was noted in 28% of cases. Slowness of movement or bradykinesia was the most commonly reported feature early in the disease, occurring in 75% of cases. Falls (60%), tremor (20%), cognitive decline (29%), speech disturbance (39%), and non specific visual symptoms (21%) were early features. Rigidity and impaired postural reflexes were found in nearly half of the patients early on in the disease course. Later in the disease, bradykinesia was present in 98% of cases, rigidity in 98% and falls in 94%. Postural reflexes were impaired in 98% of cases, after the first two years of illness, and other signs included speech disturbance (87%), dysphagia (69%), cognitive decline (74%), tremor (23%), pyramidal signs (19%) and extra-axial dystonia (26%). Cerebellar signs (1%), autonomic dysfunction (2%), and cortical sensory loss (1%) were rare. The alien limb phenomenon was not reported in any patients.

In more than half the patients, early in the disease course, examination of the eye signs was not recorded in detail, beyond a statement of "cranial nerve examination normal". However, clear documentation of supranuclear gaze palsy was present early in the disease in 38%, and pursuit or saccadic eye movements were abnormal in 44% of cases. By contrast later in the disease, 91% of cases were recorded as having supranuclear gaze palsy though the clinical examination findings were incompletely recorded in 21%.

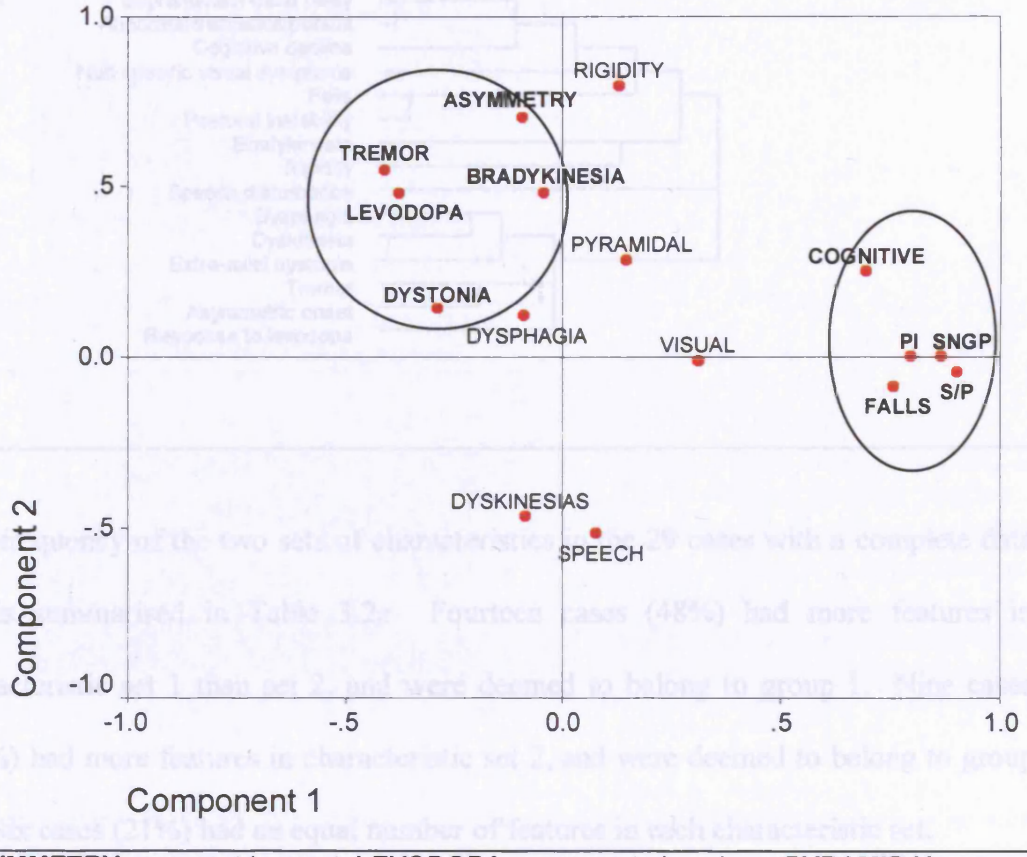
A qualitative trial of levodopa or dopamine agonist medications was undertaken in 91 of 103 cases (88%). The trial was not recorded or was not performed in the remainder. In 32% of cases there was a better than 30% improvement in symptoms coincident with the initiation of the medication. The duration of treatment, and

continued efficacy were not measured. In four cases dyskinesias developed (4% of those treated).

3.2.1.4.2. Statistical analysis

The first two components from the principal component analysis performed on data from the first two years of disease explained 42% of the variance seen in all cases. The loading of each variable on one component was plotted against its loading on the other component (Fig. 3.1). This suggested that the variables fell into two approximate groups (characteristic set 1 with high positive loadings on component 1 and characteristic set 2 with positive loadings on component 2 and negative loadings on component 1). The variables grouped together in each characteristic set were: set 1 – falls, cognitive decline, supranuclear gaze palsy, abnormal saccadic/pursuit movements, postural instability; set 2 – tremor, bradykinesia, asymmetrical onset, extra axial dystonia, response to levodopa.

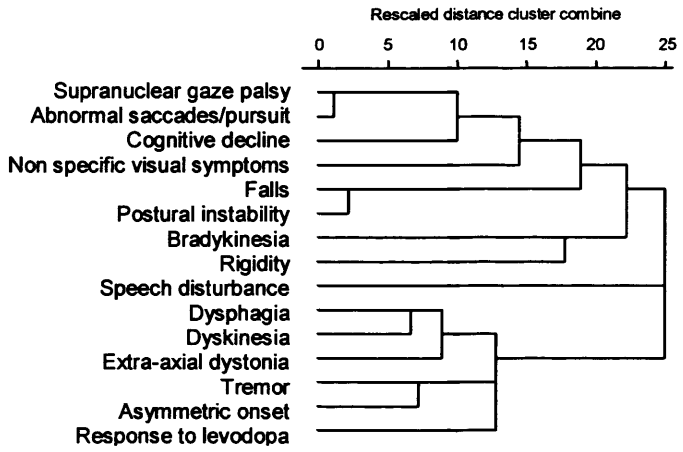
Figure 3.1. *Plot of factors for components 1 and 2, derived from factor analysis*



ASYMMETRY, asymmetric onset, **LEVODOPA**, response to levodopa; **PYRAMIDAL**, pyramidal tract signs, **COGNITIVE**, cognitive dysfunction; **VISUAL**, non-specific visual symptoms; **PI**, postural instability; **SNGP**, supranuclear gaze palsy; **S/P**, abnormal saccadic or pursuit movements; **SPEECH**, speech disturbance.

A cluster analysis of clinical variables confirmed this grouping of variables as reasonable (Fig. 3.2). This suggests that sufficient information was incorporated into the first two components of variance in order to determine groupings of characteristics between patients.

Figure 3.2 *Hierarchical cluster analysis of clinical variables: dendrogram using average linkage (between groups)*



The frequency of the two sets of characteristics in the 29 cases with a complete data set is summarised in Table 3.2a Fourteen cases (48%) had more features in characteristic set 1 than set 2, and were deemed to belong to group 1. Nine cases (31%) had more features in characteristic set 2, and were deemed to belong to group 2. Six cases (21%) had an equal number of features in each characteristic set.

The frequency of the two sets of characteristics in the 71 cases with incomplete data is shown in Table 3.2b. Five cases (7.4%) had an equal number of features in each characteristic set. Three cases were excluded from the analysis because of insufficient clinical information. The low percentage of cases that were unclassified according to this division suggests that the identified characteristics are reasonable at distinguishing two groups. The two groups share few variables from each characteristic set, suggesting that they are clinically distinct.

Table 3.2a *Cross tabulation of the number of characteristics in each of two sets exhibited by each case for initial data set.*

		Set 2					Total
		0	1	2	3	4	
Set 1	0	1	1	1	2	2	7
	1	0	2	0	0	0	2
	2	1	0	3	2	0	6
	3	0	0	0	0	1	1
	4	2	2	1	1	0	6
	5	0	6	1	0	0	7
Total		4	11	6	5	3	29

Table 3.2b *Cross tabulation for the second data set where at least one variable was missing.*

		Set 2					Total
		0	1	2	3	4	
Set 1	0	2	7	8	6	2	25
	1	1	0	0	0	0	1
	2	0	1	1	1	0	3
	3	2	10	4	2	0	18
	4	1	3	3	0	0	7
	5	3	8	2	4	0	17
Total		9	29	18	13	2	71

Set 1: falls, cognitive decline, supranuclear gaze palsy, abnormal saccades/pursuit movements, postural instability. Set 2: tremor, bradykinesia, asymmetric onset, non-axial dystonia, response to levodopa

The 66 cases that were not included in the principal components analysis, and were able to be separated into two groups with the cross tabulation, were compared according to the presence or absence of early clinical features (Table 3.3).

The co-occurrence of clinical features appeared to be similar to the cases with a complete data set, included in the principal components analysis. Early bradykinesia occurred in around three quarters of cases in both groups and did not differentiate them.

Table 3.3 *Early clinical features in cases with an incomplete data set*			
	Group 1 *% reported as present*	**Group 2** *% reported as present*	*p value*
Symptoms			
*Falls	85.7 (36/42)	0% (0/24)	<0.001[b]
*Cognitive decline	50.0 (19/38)	0 (0/22)	<0.001[b]
*Tremor	9.8 (4/41)	39.1 (9/23)	0.007[b]
Speech disturbance	32.5 (13/40)	26.1 (6/23)	0.406[b]
Dysphagia	2.7 (1/37)	4.3 (1/23)	0.624[b]
Other visual symptoms	23.1 (9/39)	4.3 (1/23)	0.051[b]
Clinical signs			
*Asymmetric onset	17.9 (7/39)	45.0 (9/20)	0.030[b]
*Bradykinesia	75.6 (31/42)	73.9 (17/23)	0.880[a]
Rigidity	39.5 (15/38)	52.5 (12/23)	0.333[a]
*Postural Instability	84.4 (27/32)	7.7 (1/13)	<0.001[b]
*Extra-axial dystonia	0 (0/38)	8.7 (2/21)	0.138[b]
*Supranuclear gaze palsy	70.0 (7/10)	0 (0/9)	0.002[b]
*Abnormal saccades/pursuits	63.6 (7/11)	0 (0/6)	0.017[b]
Pyramidal tract signs	7.9 (3/38)	4.3 (1/23)	0.513[b]
*Response to Levodopa ever	14.3 (5/35)	50.0 (11/22)	0.005[b]
* variables used in initial analysis; [a] chi-square analysis; [b] Fisher's exact test Group 1: Richardson's Syndrome, Group 2: PSP-P			

All 89 cases that could be separated into groups were compared according to profile (Table 3.4) and the late clinical features that they displayed (Table 3.5). There was a difference in sex distribution between the two groups. Men were over represented in group 1 (64.3%, 95% confidence interval (CI): 52-77%), but sex distribution was equal in group 2 (51.5%, 95% CI 33-67%). The age of disease onset was not different between the two groups.

Table 3.4 *Patient profiles*			
	Group 1	Group 2	*p value
Sex *(male, %)*	64.3	51.5	
Age at disease onset *(years)*	66.1	66.4	0.872
Age at death *(years)*	72.1	75.5	0.041
Disease duration *(years)*	5.94	9.12	<0.001
* student's t test			

Table 3.5 *Late clinical features in 89 cases separated into distinct groups*			
	Group 1 *% reported as present*	Group 2 *% reported as present*	*p value*
Symptoms			
Falls	100 (56/56)	80.6 (25/31)	0.001[a]
Cognitive decline	90.7 (49/54)	51.6 (16/31)	<0.001[a]
Tremor	13.0 (7/54)	43.8 (14/32)	0.002[b]
Speech disturbance	90.2 (46/51)	81.3 (26/32)	0.242[a]
Dysphagia	75.5 (37/49)	56.3 (18/32)	0.070[a]
Other visual symptoms	56.6 (30/53)	34.4 (11/32)	0.047[a]
Clinical signs			
Bradykinesia	98.2 (55/56)	96.9 (31/32)	0.685[a]
Rigidity	98.1 (53/54)	96.8 (30/31)	0.687[a]
Postural Instability	100 (51/51)	96 (23/24)	0.142[a]
Extra-axial dystonia	21.2 (11/51)	42.4 (13/32)	0.054[b]
Supranuclear gaze palsy	100 (50/50)	71.4 (15/21)	<0.001[a]
Pyramidal tract signs	17.0 (9/53)	15.6 (5/32)	0.561[b]
Levodopa induced dyskinesia	1.9 (1/52)	6.3 (2/32)	0.323[b]
[a] chi-square analysis; [b] Fisher's exact test Group 1: Richardson's syndrome, Group 2: PSP-P			

The ages at death and disease duration were significantly different: those in group 1 died at a younger age, and after shorter disease duration.

Falls, cognitive decline and supranuclear gaze palsy continued to be significantly associated with group 1 later in the disease ($p<0.001$). Tremor was the only clinical feature significantly associated with group 2 later in the disease though extra-axial dystonia was also more frequent in that group. There was no significant difference in

the frequency of bradykinesia, speech disturbance, postural instability, pyramidal tract signs or levodopa induced dyskinesias.

3.2.1.5. Conclusions

These data seem to confirm that two distinct clinical phenotypes exist in patients with pathologically proven PSP. One is characterised by early falls, early cognitive dysfunction, abnormalities of gaze and postural instability, and the other by asymmetric onset, tremor, early bradykinesia, non-axial dystonia and a response to levodopa medications. The clinical characteristics present in this first group are similar to those first described in PSP by Richardson (Richardson *et al.*, 1963; Steele *et al.*, 1964), whereas the clinical features in the second group more closely resemble PD. Phenotype one has a shorter duration of disease and the female to male ratio is 1:1.8, whereas the sex distribution is equal in phenotype two (see Table 3.4).

Because of ascertainment bias the proportions of patients presenting with these phenotypes in this study may not reflect the proportions of these subtypes in the community (Maraganore *et al.*, 1999). However the frequency of falls (94%), bradykinesia (98%), speech disturbance (87%) and dysphagia (69%) in the whole group was consistent with data reported in other clinical and clinicopathological studies (Maher and Lees, 1986; Nath *et al.*, 2003; Verny *et al.*, 1996b). In contrast to some studies there was a higher incidence of levodopa responsiveness (32%), extra-axial dystonia (26%) and a longer duration of disease (Maher and Lees, 1986; Nath *et al.*, 2003). The inclusion of only pathologically proven cases of PSP partially accounts for this, and is a strength of this study, allowing us to more clearly identify atypical clinical features in cases that would have been automatically excluded from clinical reports relying on PSP clinical diagnostic criteria.

Previously, attempts have been made to define clinical subgroups in cohorts of PSP (Birdi *et al.*, 2002; Gibb *et al.*, 2004; Maher and Lees, 1986; Morris *et al.*, 2002; Nath *et al.*, 2003) and to apply this classification in small pathologically proven series (Bergeron *et al.*, 1997; Braak *et al.*, 1992; Daniel *et al.*, 1995; Verny *et al.*, 1996b). Daniel and co-workers found that in a subgroup without supranuclear gaze palsy women were over represented, age at onset was later and duration of disease was longer when compared to the group with supranuclear gaze palsy, where men were over represented (Daniel *et al.*, 1995). No quantitative or qualitative pathological differences were distinguished, although a subsequent study on the same material showed greater neuronal loss in the nucleus interpositus in those with a supranuclear gaze palsy (Revesz *et al.*, 1996). Birdi and collaborators reported that patients with gait or balance difficulties at onset were less likely to improve on levodopa therapy, more likely to develop supranuclear gaze palsy and had a shorter duration of disease (Birdi *et al.*, 2002). Morris and colleagues classified pathologically confirmed cases of PSP according to the diagnosis of PSP in life and the retrospective application of diagnostic criteria for PSP (Morris *et al.*, 2002). All cases in their "typical PSP" subgroup had a diagnosis of PSP in life, all retrospectively satisfied the diagnostic criteria, and such cases were more likely to have the PSP susceptibility tau haplotype (H1,H1). In a series where only 60% of cases had pathological confirmation, Nath and co-workers found that the subgroup with early falls and bulbar dysfunction had a shorter survival, and those with a diagnosis of probable PSP according to the National Institute for Neurological Diseases and Stroke – Society for PSP (NINDS-SPSP) criteria had a worse prognosis (Nath *et al.*, 2003). The variations in definition between these studies make it difficult to compare clinical groups and even more difficult to apply clinical characteristics objectively to pathological cohorts. By applying the

clinical distinctions identified in this study, comparisons can be made between more homogenous clinical groups potentially enabling a better understanding of the spectrum of pathological changes in PSP.

A number of authors have found pathologic differences in cases of clinically "typical" and "atypical" PSP, without applying a standardised system for clinical classification. In one report two subgroups were identified that could be divided on clinical and pathological grounds (Verny *et al.,* 1996b). A group with mild involvement of the pedunculopontine nucleus invariably had mild cortical involvement and a number of "unusual" clinical signs including absence of oculomotor palsy and axial rigidity, severe early dementia and rest tremor. The subcortical structures involved had a preferentially 'pallido-luyso-nigral' distribution. The second group all had typical clinical signs with more severe involvement of the pedunculopontine nucleus and cortex, and more widespread involvement of subcortical structures. The sample size, however, was too small to come to definitive conclusions relating to differences. In other reports patients with early severe dementia were found to have more cortical pathology (Bigio *et al.*, 1999; Braak *et al.,* 1992) , and pathological changes in the nucleus raphe interpositus which correlate with supranuclear gaze palsy in PSP have been reported (Revesz *et al.,* 1996). The variation of pathological involvement in PSP, and the identification of some clinical correlation with it, supports the notion that different clinical groups exist and differences in early pathological involvement of the striatum, pallidum or subthalamus may account for the observed marked clinical differences between these different presentations of PSP.

.

No cases exhibiting any of the NINDS-SPSP mandatory exclusion criteria were identified (Litvan *et al.,* 1996a). However, according to the data some features

included in the NINDS-SPSP "supportive criteria" may be misleading. Early speech disturbance was found in 39% of cases, and 32% had a modest or good response to levodopa (PDSBB grade 2 or 3), though an excellent response to levodopa (PDSBB grade 4) was never recorded. While an asymmetrical onset was a feature in 28%, persistent markedly asymmetrical Parkinsonism was not seen.

With the results of this preliminary data analysis in mind, the collection of clinical data in this imaging study was prospectively planned. The principal aim was to look at rates of brain atrophy, but in order to relate brain atrophy to the clinical features and determine whether the phenotypic variability has a clinical correlate, an accurate and standard method of clinical data collection needed to be used. The datasheet used to retrospectively collect data in the post mortem study was altered for use in this prospective study. Clinical rating scales including the UPDRS were used so that parameters such as bradykinesia, rigidity, cognitive decline, mobility and postural stability could be quantified.

Full details of clinical data collection, is provided in the following sections and in chapter 4. The clinical exclusion criteria for PSP were adhered to strictly as the results of this preliminary study indicate that they are helpful in preventing a false positive clinical diagnosis of PSP.

3.3. The clinical diagnostic criteria

The current accepted clinical diagnostic criteria for PSP (Litvan *et al.*, 2003), MSA (Gilman *et al.*, 1999) and PD (Gibb and Lees, 1988) are outlined below (Tables 3.6a-c).

Table 3.6a Diagnostic criteria for PSP

PSP

Inclusion criteria:	**Supportive criteria:**	**Exclusion criteria:**
For clinically probable and definite: Gradually progressive disorder with age at onset at 40 or later; **Clinically possible PSP:** Gradually progressive disorder >12 months duration, onset over 40 yrs, tendency to fall within 1 year of onset. **Clinically probable PSP:** Either vertical supranuclear palsy or both slowing of vertical saccades & postural instability with falls < 1 yr disease onset **Clinically definite PSP:** Vertical supranuclear palsy and prominent postural instability with falls/tendency to falls, within first year of disease onset **Definite PSP:** All criteria for possible or probable PSP are met and histopathologic confirmation at autopsy	Symmetric akinesia or rigidity, proximal more than distal; abnormal neck posture, especially retrocollis; poor or absent response of parkinsonism to levodopa; early dysphagia & dysarthria; early onset of cognitive impairment including > 2 of: apathy, impairment in abstract thought, decreased verbal fluency, utilization or imitation behaviour, or frontal release signs	Recent history of encephalitis; alien limb syndrome; cortical sensory deficits; focal frontal or temporoparietal atrophy; hallucinations or delusions unrelated to dopaminergic therapy; cortical dementia of Alzheimer type; prominent, early cerebellar symptoms or unexplained dysautonomia; or evidence of other diseases that could explain the clinical features

Table 3.6b Diagnostic criteria for MSA

MSA

Autonomic and urinary dysfunction	Orthostatic hypotension (by 20 mm Hg systolic or 10 mm Hg diastolic); urinary incontinence or incomplete bladder emptying	Orthostatic fall in blood pressure (by 30 mm Hg systolic or 15 mm Hg diastolic) and/or urinary incontinence (persistent, involuntary partial or total bladder emptying, accompanied by erectile dysfunction in men)
Parkinsonism	B, R, I, and T	1 of 3 (R, I, and T) and B
Cerebellar dysfunction	Gait ataxia; ataxic dysarthria; limb ataxia; sustained gaze-evoked nystagmus	Gait ataxia plus at least one other feature
Corticospinal tract dysfunction	Extensor plantar responses with hyperreflexia	No corticospinal tract features are used in defining the diagnosis of MSA

B, bradykinesia; R, rigidity; I, postural instability; T, tremor

Table 3.6c Diagnostic criteria for PD

PD

Inclusion criteria	**Supportive criteria**	**Exclusion criteria**
Bradykinesia (slowness of initiation of voluntary movement with progressive reduction in speed and amplitude of repetitive actions) And at least one of the following: Muscular rigidity. 4-6 Hz rest tremor. Postural instability not caused by primary visual, vestibular, cerebellar, or proprioceptive dysfunction.	*(Three or more required for diagnosis of definite PD)* Unilateral onset. Rest tremor present. Progressive disorder. Persistent asymmetry affecting side of onset most. Excellent response (70-100%) to levodopa. Severe levodopa-induced chorea. Levodopa response for 5 yr or more. Clinical course of 10 yr or more.	History of repeated strokes with stepwise progression of parkinsonian features History of repeated head injury History of definite encephalitis Oculogyric crises Neuroleptic treatment at onset of symptoms More than one affected relative Sustained remission Strictly unilateral features after 3 yr Supranuclear gaze palsy Cerebellar signs Early severe autonomic involvement Early severe dementia with disturbances of memory, language, and praxis Babinski sign Presence of cerebral tumour or communicating hydrocephalus on CT scan Negative response to large doses of levodopa (if malabsorption excluded) MPTP exposure

From the most recent update on research clinical diagnostic criteria (Litvan *et al.*, 2003), patients are considered to have clinically possible PSP if they suffer from a gradually progressive disorder of more than 12 months duration, with onset over 40 years of age, and with a tendency to fall within the first year of disease onset, in the absence of defined exclusion criteria (Table 3.6a). There should be no clinical features suggestive of Creutzfeldt-Jakob disease or any other identifiable cause for their postural instability.

3.4. Planning the neuropsychological assessments

A number of neuro-cognitive models have been proposed for "frontal" cognition. These are inevitably incomplete. These are associated with structure function correlations which are also inadequate. Despite these caveats, the following section may be useful as background to the results described in later sections.

3.4.1. Frontal sub-cortical circuit syndromes

A series of parallel frontal-subcortical circuits that link specific regions of the frontal cortex, important in human behaviour, to the striatum, globus pallidus and the thalamus have been defined (figure 3.3) (Alexander *et al.*, 1986).

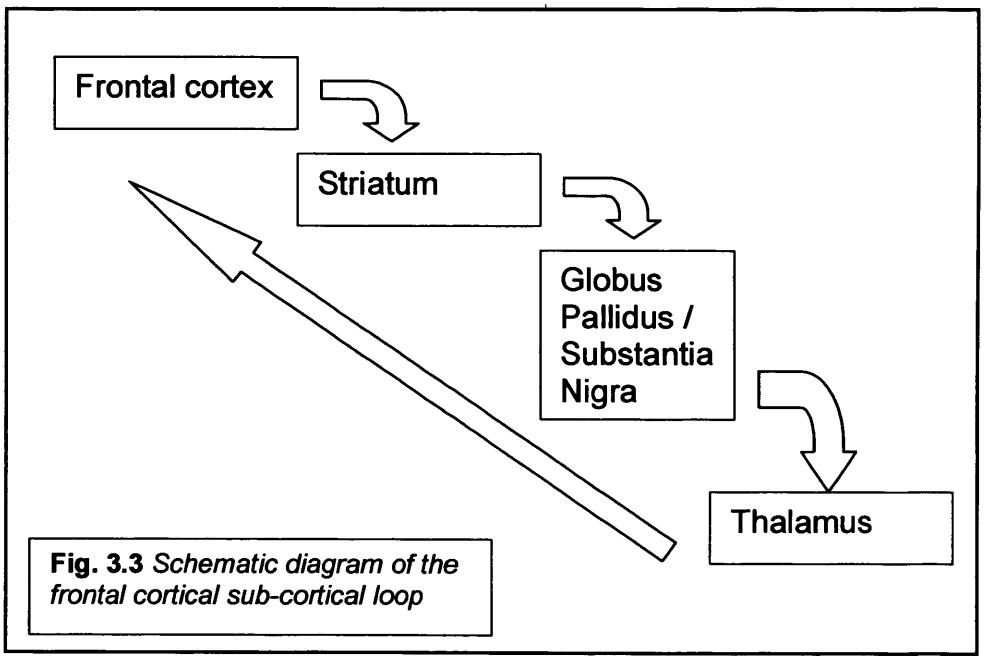


Fig. 3.3 *Schematic diagram of the frontal cortical sub-cortical loop*

Five circuits named according to function or site of origin exist.

1. The motor circuit (originating in the supplementary motor area).
2. The oculomotor circuit (originating in the frontal eye fields).
3. The dorsolateral prefrontal circuit (originating in Brodmann area 9+10).
4. The orbital-frontal circuit (originating in Brodmann area 10+11).
5. The anterior cingulate circuit (originating in Brodmann area 24).

More recently, it has been suggested that the original five-circuit scheme should be broadened to include seven general categories of circuits: skeletomotor; oculomotor; dorsolateral prefrontal; lateral orbitofrontal; medial orbitofrontal; cingulate and inferotemporal-posterior parietal (Middleton and Strick, 2000).

Within each circuit, there is a direct and an indirect pathway. These modulate input to the thalamus in response to different stimuli.

Fibres in each circuit originating from the frontal lobe are mediated by excitatory glutaminergic neuro-transmission. These projections are mainly to the striatum. The dorsal striatum has two separate cytochemical structures, the striosomes and the matrix. The striosomes mature earlier; have a lower dopamine concentration and a higher concentration of limbic associated membrane protein. The striosomes receive input mainly from the orbitofrontal, temporal and insular cortex and the ventral substantia nigra (pars compacta). Output is mainly to the medial substantia nigra. The matrix receives input from the sensorimotor cortex and the parietal cortex as well as the cingulate gyrus and output is mainly to the globus pallidus and the substantia nigra.

3.4.1.1. Dorsolateral prefrontal syndrome

Patients with damage to the dorsolateral prefrontal circuit, are concrete and perseverative with slow and impaired reasoning and poor mental flexibility. They have an inability to maintain and to redirect attention. Typically they perform poorly on tests of attention and set shifting such as the Wisconsin Card Sorting Test. They have difficulties with verbal fluency and alternating motor sequences.

3.4.1.2. Orbitofrontal syndrome

The orbitofrontal circuit connects frontal monitoring systems to the limbic system. Dysfunction here is characterised by behavioural disinhibition and emotional lability. Patients are irritable and may have explosive aggressive outbursts. Inappropriate responses to social cues, lack of interpersonal sensitivity, improper sexual remarks and inappropriate jocularity are other clinical features.

3.4.1.3. Anterior cingulate syndrome

The anterior cingulate circuit mediates motivated behaviour and dysfunction in this area is reflected by apathy. A bilateral anterior cingulate lesion results in akinetic mutism, which has the features of prominent apathy, an indifference to pain, thirst and hunger, inertia, a lack of motor initiative, and a lack of verbalization and response to commands. Abulia is a less severe form of this psychomotor retardation.

The major deficit of patients with damaged anterior cingulate circuit structures is response inhibition on go/no-go tests.

The cognitive deficit seen in PSP has features consistent with damage to all of these frontal sub-cortical circuits.

3.4.2. Review of the Neuropsychological profile in PSP

In the seminal description of PSP, emphasis was placed upon the distinctive clinical features of early postural instability, a supranuclear gaze palsy predominantly affecting vertical eye movements, axial rigidity, retrocollis and symmetrical Parkinsonism (Richardson *et al.,* 1963; Steele *et al.,* 1964). Cognitive dysfunction was also noted in seven of the nine patients described in 1964. In one case this was limited to mild bradyphrenia but in the others, emotional lability, increased irritability and personality change were present. While these changes were often reported early, they tended to remain mild and in only two cases became severe. Only one of the nine cases had early severe dementia.

Since the original description, the neuropsychological features of PSP have been extensively reviewed (Albert *et al.,* 1974; Grafman *et al.*, 1995; Litvan *et al.*, 1989; Litvan, 1994; Litvan *et al.*, 1996d; Maher *et al.*, 1985; Robbins *et al.*, 1994). Albert *et al* further expanded the cognitive profile of PSP with a study documenting the neuropsychological findings in five of their own cases of PSP and 42 other cases from the medical literature where sufficient information was available (Albert *et al.,* 1974). They described a constellation of symptoms including slowed thought processes, emotional lability, personality change and an impaired ability to manipulate acquired knowledge, which they referred to as "*subcortical dementia*".

This syndrome of subcortical dementia has been claimed to be distinct from cortical dementia in its clinical manifestation and in its anatomico-pathological correlations.

Sub-cortical dementia describes a combination of profound slowing of cognition, memory disturbance and executive dysfunction as well as changes in personality and affect (Albert *et al.,* 1974). It has been argued however that neuropsychological profiles of cortical and sub-cortical dementia are not sufficiently distinct (Brown and

Marsden, 1988) and that cortical cognitive deficits may occur with subcortical lesions (Hughes *et al.*, 1993).

In Alzheimer disease (AD), typical clinical findings include dyscalculia, dysphasia, dyspraxia and agnosia. Accelerated forgetting is also said to be striking. These are said to result from cortical dysfunction. Apathy and irritability along with depression are less common in AD than they are in Parkinson's disease (PD) (Aarsland *et al.*, 1999)

Executive dysfunction is one of the main features of sub-cortical dementia and consists of difficulty with verbal fluency, set shifting, categorisation and planning. There is some evidence that this is may also be impaired in early AD (Kopelman, 1991). In sub-cortical dementia, learning impairment can be partially corrected by providing richer or more salient cues to encourage learning and to promote recognition.

None of the PSP patients Albert *et al* reviewed in 1974 had evidence of aphasia, agnosia or apraxia (Albert *et al.,* 1974), although mild apraxia may occur in PSP (Leiguarda *et al.*, 1997; Pharr *et al.*, 1999; Pharr *et al.*, 2001) and aphasia has been rarely reported (Esmonde *et al.,* 1996). A quantitative study of the cognitive profile in PSP found that bradyphrenia was evident but short-term memory was unaffected (Litvan, 1994). In PSP, difficulties with concept formation and with shifting established conceptual sets were common, as were slowed psychomotor responses. Attention was impaired and this was related to disease duration.

Subcortical lesions have been well documented to cause a specific pattern of dementia. Symptoms of mental impairment such as inattention, disorientation, confusion and emotional disturbances dominated the clinical picture in an early clinico-pathological study of tumours of the thalamus (Smyth and Stern, 1938).

Various other reports of dementia in the absence of cortical pathology appear in the literature. These variously describe, some form of memory defect, a general slowing of intellectual activity, alterations of personality (typically, inertia or apathy, occasionally with episodic irritability) and an impaired ability to manipulate acquired knowledge.

Visual attention is affected in PSP, and patients have difficulty orientating their attention when compared to controls. The severity of this impairment correlates well with the course of the disease. Hence, all tasks requiring visual scanning may result in impaired performance in PSP (Kimura *et al.*, 1981). This may mean that some aspects of the pattern or severity of cognitive impairment described in PSP arise as a consequence of the psychological test scores being influenced by defective visual scanning and are not an integral feature of the disease. More recent testing using techniques not dependent on visual scanning support the presence of a clear pattern of dementia but as stated this is not typically early or severe. Defective visual scanning needs to be taken into account in any test battery for PSP.

Patients with PSP have a slowed speed of information processing. This is in part due to slowed motor response times, but sensory and cognitive information processing are also slowed and subject to degradation. Working memory (short term memory stores), from which information is transferred to more permanent knowledge stores, is unaffected in PSP. However, several aspects of episodic (long-term) memory appear to be impaired. Verbal learning with the Rey Auditory Verbal learning test has shown that patients with PSP recalled fewer words than controls. Cueing improves performance (Pillon *et al.*, 1994) suggesting a difficulty in accessing stored information. This contrasts with the recognition and recall deficits with hippocampal and cortical impairments. Strategic encoding and retrieval procedures subserved by

the fronto-striatal system seem particularly compromised in the "sub-cortical" dementia of PSP.

Executive functions are impaired in PSP and this is thought to be due to deafferentation between the basal ganglia and the prefrontal cortex. The prefrontal cortex has been described as supporting "executive function" including working memory, reasoning, problem solving, conceptualisation, planning and social cognition. PSP patients have problems forming concepts and interpreting proverbs. They make more errors on the Wisconsin Card Sorting test and the majority of these are of the perseverative type. They also have difficulty ordering events in the correct sequence as shown by poor performance on the picture arrangement subtest of the Wechsler Adult Intelligence Scale (WAIS). Word and figural fluency tests both result in perseverative responses in PSP.

There is little information available regarding the mood state and personality change of PSP patients following the onset of their disorder.

In summary, changes in executive function and information processing appear early in the course of PSP and are relatively severe. Both of these are characteristic of PSP and help to differentiate it from other dementias. Orientation and episodic memory although impaired, are less severely affected. The majority of these cognitive defects are attributed to frontal deafferentation secondary to sub-cortical lesions.

The executive dysfunction in PSP is also to a lesser extent, a feature in other akinetic-rigid syndromes such as MSA and PD.

Cognition and emotion have been compared in different akinetic rigid disorders. The risk of developing dementia in PD is at least 3 times greater than in age-matched subjects. A recent study, including 224 patients with PD, with a mean age of 73.4

years and a mean disease duration of 9.2 years suggested that, based on a clinicians assessment and scores on three cognitive rating scales, the eight year prevalence of dementia in PD is around 78% (Aarsland *et al.*, 2003). In PD, visuospatial, executive function and memory are impaired in the absence of genuine amnesia or impairment of instrumental activities resulting in aphasia, apraxia or agnosia. Mild executive dysfunction is observed even at the early stages of PD when lesions are probably restricted to the brainstem structures supporting the hypothesis that they derive from sub-cortical pathology of the disease. Dementia in PD is associated with depression, institutionalization, older age at onset, and atypical neurologic features, such as early occurrence of autonomic failure, symmetrical disease presentation, and only moderate response to a dopaminergic drug (Aarsland *et al.*, 1996).

Compared with PD, executive functions in PSP are much more severely affected. Although the cortex is involved in PSP, the severe frontal syndrome observed in patients with PSP seems to be as a consequence of the loss of specific afferents to the pre-frontal cortex from sub-cortical regions.

MSA is classically thought of as having normal cognitive and affective status with a mild impairment of memory and executive function, selective for frontal lobe involvement. This is qualitatively different from PD and PSP. As well as being severely affected in PSP, there is involvement to a lesser extent of the frontal functions in the cognitive changes associated with PD, however, patients with PSP tend to have significantly lower scores in frontal behavioural tests than PD or MSA at all levels of cognitive impairment. In PD there is slow verbal production and formulation of answers. There is a decline in attention, an incapacity for abstraction and criticism and impaired judgement. In addition, visuo-spatial function, executive function and memory are all affected in the absence of genuine amnesia or

impairment of instrumental activities such as aphasia, apraxia or agnosia. These features are related to the sub-cortical pathology of the disease. In PD, the biggest problems arise from difficulties with executive function involving problem solving and planning. Generally speaking, the PD patients with deteriorating cognition tend to be older, have greater motor impairment and more psychotic symptoms and hallucinations (Aarsland *et al.*, 1996).

Soliveri et al (Soliveri *et al.*, 2000) followed up 23 patients with MSA-P, 21 patients with PSP and 18 patients with PD. The patients were matched for disease severity, according to the Hoehn and Yahr scale and the activities of daily living (ADL) on the UPDRS. The rates of deterioration in most cognitive tests did not differ greatly between the groups, but the relatively short follow up period (21 months) and a high dropout rate (40%) limits definitive conclusions.

3.4.3. Choice of neuropsychological test battery

The neuropsychological test battery for this thesis was chosen in order to attempt to optimise the distinction between PSP, MSA-P and PD. Ability of the tests to detect change in cognition over time was also given consideration, but given that a relatively short interval between assessments was planned, this was felt to be of secondary importance. An experienced neuropsychologist was involved from the outset. Tests of more general cognitive function were also included in order to enable assessment of how for example memory impairment might affect other cognitive test scores.

Each patient's level of education was noted. This was scored in terms of years of education with each year of schooling after the age of 5 counted. As an example, an individual leaving school at age 16, would have 11 years of education, whereas, a

person continuing to a university degree would have 11 plus 2 (schooling from age 16-18), plus 3 (university degree) years, totalling 16 years of education.

The pre-morbid level of intellectual functioning was estimated using the National Adult Reading Test (NART) (Nelson, 1982). The neuropsychological test battery included: the Mattis dementia rating scale-2 (which consists of the following subtests: attention, verbal and non-verbal memory, motor initiation, construction and conceptual ability) with a maximum score of 144 and a cut off suggesting cognitive dysfunction of 130 (Mattis S., 1973); current Verbal IQ based on completion of the Vocabulary, Similarities and Digit Span subtests of the Wechsler Adult Intelligence Scale-revised (WAIS-R) (Weschler, 1981); the Rey Auditory Verbal Learning Test (RAVLT) (Rey A, 1958) ; the Short Recognition Memory for Faces (RMF) (Warrington, 1984); the Reitan Trail Making Test (TMT) which included Trail A, a simple measure of behavioural regulation, sequencing and motor speed and Trail B, in which the subject alternates between sequences of numbers and letters necessitating shifting of mental set. For each version, the time to complete the test is recorded. The difference score of the two versions has been used as an index of behavioural regulation and "cognitive speed" independent of motor speed (Reitan, 1958). In addition, patients were asked to carry out the Wisconsin Card Sorting Test (WCST) (Nelson, 1976); the semantic, phonemic and alternating semantic versions of the verbal fluency test (Benton, 1968) and the paced auditory serial addition test (PASAT) (Gronwall and Wrightson, 1981). At the end of the session; the Apathy Scale (Marin, 1990), the Beck Anxiety and Depression Inventories (Beck *et al.*, 1961; Beck and Steer, 1990), were completed by all patients (Table 3.7).

While PSP and MSA subjects typically may only have minimal levodopa responsiveness, some patients do report subjective improvement with levodopa. All

patients in this study were assessed in the "on" state with regard to their levodopa response.

Table 3.7 *Neuropsychological test battery completed by each subject*

Mattis Dementia Rating Scale (Mattis DRS)
Rey Auditory Verbal Learning Test (Rey AVLT)
WAIS-R
NART
Wisconsin Card Sorting Test
Reitan Trail Making Test A and B
Recognition Memory Test
Paced Auditory Serial Addition Test
Verbal Fluency Test
Beck Anxiety Inventory
Beck Depression Inventory
Apathy Questionnaire

3.5. Planning the imaging

A preliminary, retrospective review of serial volumetric MRI scans acquired on patients referred to the cognitive disorders clinic and a specialised movement disorders clinic at the National Hospital for Neurology and Neurosurgery, Queen Square, was undertaken. The purpose of this was to identify individuals with PSP and MSA in whom serial imaging had been undertaken, and in whom the diagnosis had been pathologically confirmed.

3.5.1. Methods

A retrospective review of the database from the specialist cognitive disorders clinic and a specialist movement disorders clinic was undertaken. Three cases with pathologically confirmed PSP and one with pathologically confirmed MSA were identified, in whom serial MRI had been undertaken on the same scanner, during life. In two of the PSP cases, movement artefact on the second scan rendered them unsuitable for image registration. Hence the clinical histories and MRI appearances in

one PSP patient and one MSA patient were reviewed. The post mortem histopathological features in the two cases were compared to the imaging findings.

3.5.2. Case reports

3.5.2.1. PSP Case

A 59 year-old male presented with a three-month history of non-specific symptoms including fear that he may be developing dementia. His son commented that his driving skills had deteriorated. Past history included hypertension controlled with bendrofluazide and enalapril, and he drank 60 units of alcohol a week. Examination by a neurologist revealed no focal neurological signs. Over the next six months his balance deteriorated and there was a change in his speech.

He was seen at the cognitive disorders clinic at The National Hospital for Neurology and Neurosurgery. By this time he had had several falls and he felt "clumsier". His speech was slurred and he had difficulty in adjusting focus from near to distant objects. He was admitted to hospital where he was noted to have absent convergence, supranuclear vertical gaze palsy and slow vertical saccades. Smooth pursuit was impaired and optokinetic nystagmus was sluggish with eye deviation in the direction of the slow phase. He had a cautious hesitant gait, postural instability and bradykinesia but minimal limb rigidity. There was no improvement an hour post Madopar-250mg. Autonomic function testing was normal. Neuropsychological testing showed that he had significant cognitive impairment particularly in frontal domains.

A diagnosis of PSP was made. His condition progressed relentlessly and he developed axial rigidity, frequent falls, cervical dystonia, a worsening dysarthria and dysphagia. He died 5 years after disease onset, 19 months after his second scan.

Cross sectional imaging during life was reported simply as showing global atrophy.

3.5.2.2. MSA Case

A 60 year-old right-handed man was referred with a three year history of increasing unsteadiness. At the time of initial assessment, he had limited mobility and difficulty getting in and out of his car. He also reported increasingly slurred speech, problems swallowing and difficulty with dextrous movements. He was also impotent and had nocturia and alteration in bowel habit. His sleep was disturbed by laboured breathing and he occasionally shouted and struck out during the night. The patient and his family reported that his memory had been deteriorating for approximately 1 year, and that he had become increasingly irritable and aggressive. Past medical history included lumbar discectomy and tonsillectomy. He was a nonsmoker who drank very little; he only took Gaviscon, and there was no relevant family history.

On examination he scored 21/30 on the Mini-Mental State Examination. He had a mask-like facies and bradykinesia with slurred monotonous speech. Examination of eye movements revealed jerky pursuit and nystagmus on lateral gaze. There was cogwheel rigidity and bradykinesia in the limbs. The leg reflexes were brisk, and the plantar responses extensor. There was finger-nose dysmetria, dysdiadochokinesis and heel-shin ataxia, and diminution of vibration sense at the ankles. The gait was slow and broad-based. There was no postural hypotension.

The following investigations were normal or negative: full blood count, urea and electrolytes, liver function, thyroid function, B12, folate, ESR, VDRL, chest X-ray, antineuronal antibodies and genetic testing for SCAs 1, 2, and 3. A CSF examination was unremarkable apart from a slightly elevated protein (0.7 g/l). Autoantibody testing revealed a positive anti-nuclear antibody (ANA) titre at 1/320, and strongly positive anti-axonal antibodies of uncertain significance. Formal autonomic function

testing confirmed no orthostatic hypotension and a normal heart rate response, but absent pressor responses and a partially blocked valsalva manoeuvre, thought to indicate possible early autonomic failure. A sphincter electromyogram (EMG) was abnormal, and suggestive of a diagnosis of MSA. A first MRI was carried out in December 1996, and the patient underwent neuropsychological testing. A diagnosis of probable MSA was made.

Over the following year, the patient continued to decline. He fell more frequently and experienced increasing slurring of speech, with worsening dysphagia, urinary and sexual dysfunction, and memory impairment. Around this time, he required permanent urethral catheterisation, having gone into urinary retention. By 1998, he was unable to walk unaided, and was dependent for all activities of daily living. Examination at this time revealed square-wave jerks, gaze-evoked nystagmus and impaired tongue movements. There was marked generalised axial and limb rigidity with bradykinesia, but no tremor. The reflexes were now all brisk and he had bilateral Babinski signs. An ear, nose, and throat examination revealed normal vocal cords, and there was no response to a formal challenge with Co-beneldopa (Madopar 250; Roche, SA). A second MRI and repeat neuropsychological testing were carried out. The clinical features were again thought to be consistent with a diagnosis of MSA. There was continued decline over the following months, and the patient died 1 year later (February 1999). A postmortem examination was carried out.

Neuropsychological testing was obtained on two occasions, approximately 1 year apart. A comprehensive battery was used, including estimates of intelligence using the National Adult Reading Test (NART) and Wechsler Abbreviated Scale of Intelligence-Revised (WAIS-R) scores. Memory was assessed using the Recognition

Memory Tests (RMT), nominal skills using the Graded Difficulty Naming Test (GNT), and perceptual functioning was assessed using the Fragmented Letters and Cube Analysis of the Visual Object and Space Perception Battery (VOSP). Frontal lobe function was assessed using the Weigl colour-form sorting test and tests of verbal fluency. At the second assessment, additional tests sensitive to speed and attention (months of the year backwards and counting back from 30 to 1) were carried out.

The first MR scan revealed marked atrophy of the cerebellum and moderate atrophy of the pons and middle cerebellar peduncles. By the second scan, 14 months later, there had been progression in the infratentorial atrophy, particularly of the pons and middle cerebellar peduncles, the former demonstrating the "hot-cross bun" configuration seen in MSA.

3.5.2.3. MRI protocol

T1-weighted volumetric MR scans were acquired at two points, (14 months apart in the MSA case and 18 months apart in the PSP case) on the same 1.5-T GE Signa Unit (General Electric, Milwaukee, WI) using a spoiled gradient-echo technique (256 × 192 matrix, FOV 20 × 20 cm, TR/TE/NEX/FLIP = 15.8/4.2/1/20°). This yielded 124 contiguous 1.5-mm thick slices, which were transferred to a Sun Enterprise 250 workstation (Sun Microsystems, Mountain View, CA) for analysis. T2-weighted and proton-density scans were acquired at each time point. An experienced neuroradiologist reported all MRI scans. Image processing was carried out using the MIDAS software tool (Freeborough *et al.*, 1997). A semiautomated technique, using intensity thresholding and a series of erosions and dilations, was applied to delineate brain tissue.

Using the rigidly aligned repeat scan as a starting point, the model generates a voxel-level deformation field to match the scans accurately. The determinant of the Jacobian matrix from the fluid model gives a voxel-by-voxel measure of volume change, and voxel compression maps are created using a colour overlay to signify this volume change (details are given in chapter 2.7.3 and chapter 10) (Crum *et al.*, 2001).

3.5.3. Results

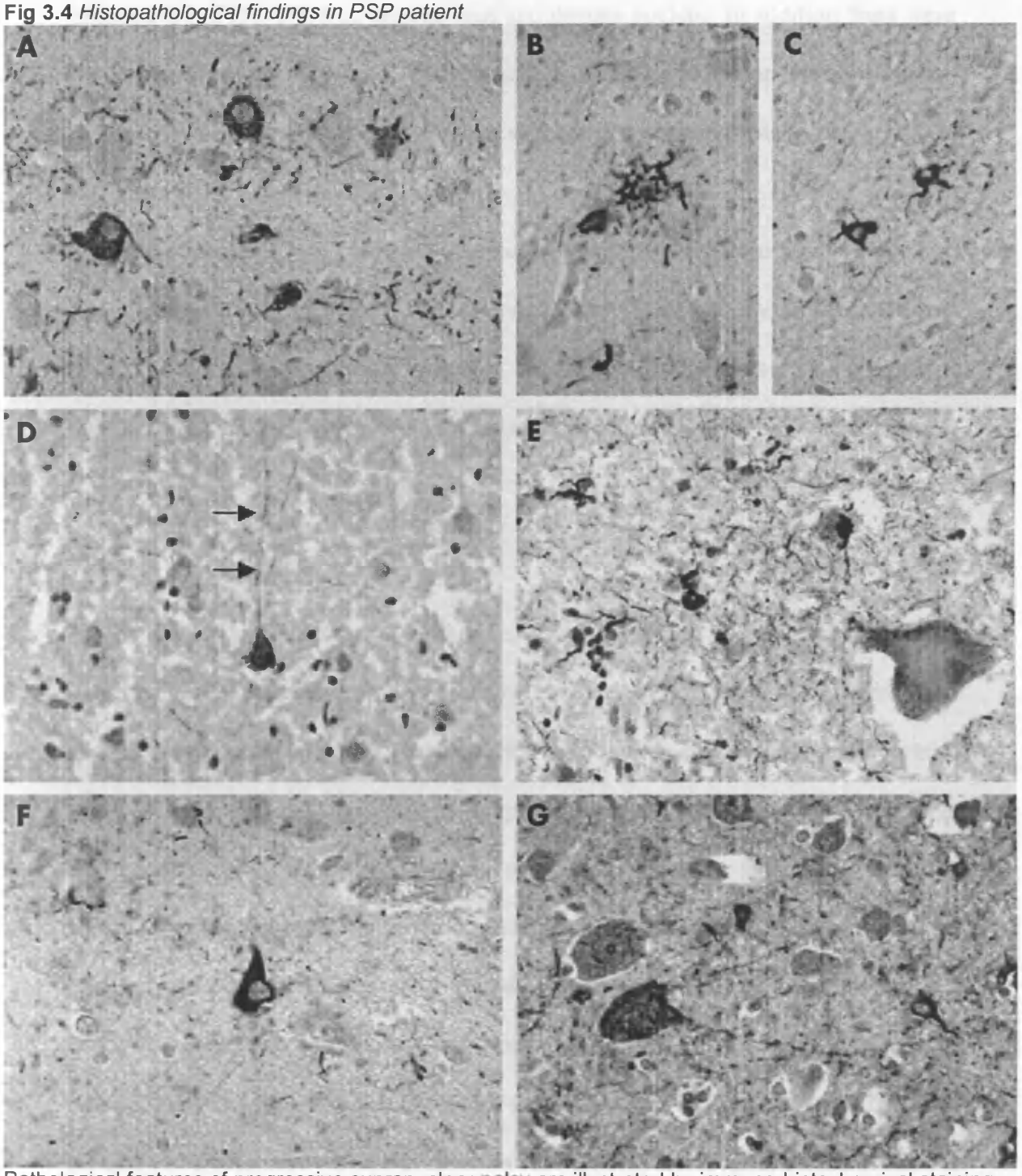
3.5.3.1.Neuropathology

3.5.3.1.1. PSP

The brain (1,529g) was bisected and the right half, flash frozen. Histological examination of the fixed tissue used routine stains and immunohistochemistry for tau (AT8 1:600, Autogen Bioclear, UK), ubiquitin (1:300, Dako, UK) and α-synuclein (1:2000, Autogen Bioclear, UK). Frozen sections (7μm) from the right anterior and posterior frontal lobe were cut and stained using tau immunohistochemistry.

Coronal slices from the formalin fixed left half brain, showed moderate dilatation of the lateral ventricle with atrophy of the globus pallidus (GP) and subthalamus (STN), which was discoloured. The substantia nigra (SN) and locus ceruleus (LC) were pale. There was severe neuronal loss with free pigment and gliosis in the SN. Moderate cell loss and gliosis was evident in the STN and LC.

Fig 3.4 *Histopathological findings in PSP patient*



Pathological features of progressive supranuclear palsy are illustrated by immuno-histochemical staining for tau. (A) Neurofibrilliary tangles and neuropil threads are present in the pontine base. (B) Tufted astrocytes are demonstrated in the anterior frontal cortex and (C) oligodendroglial inclusions in the cerebellar white matter. Comparison of the anterior frontal (D, F) and motor cortices (E, G) in each hemisphere shows more numerous tau positive structures including glial inclusions and neuropil threads in the motor cortex than in the corresponding anterior frontal cortex (arrows on D pointing to the apical dendrite of a tau positive neurone). D, E show frozen sections from the right half brain, while F, G illustrate paraffin embedded tissue from the left hemisphere (magnification x40).

Immunohistochemical staining for tau demonstrated NFTs and NTs in the neocortex, striatum, GP, STN, SN, LC, pontine base and dentate nucleus. In addition there were tau immunoreactive oligodendroglial inclusions particularly evident in the striatum, GP and cerebellar white matter. Tufted astrocytes were most easily found in the frontal cortex and putamen (Fig 3.4 A-C).

In both hemispheres, the motor cortex was more severely affected than the anterior frontal cortex (Fig 3.4 D-G). It was not possible to determine any difference in the amount of tau pathology between the right and the left frontal lobes due to processing differences and no absolute neuronal counts were obtained.

3.5.3.1.2. MSA

At postmortem examination the brain was of normal weight (1276 g), although the brainstem and cerebellum were light by comparison (122 g). On macroscopic examination, the most notable abnormalities were the small size of the basis pontis, pallor of the substantia nigra, and atrophy of the cerebellum. There was very mild sulcal widening over the frontal poles and slight dilatation of the lateral ventricles. On microscopy, the temporal neocortex showed some slight superficial spongiosis. The hippocampus was well preserved; amyloid plaques were absent and only occasional tangle-like structures were seen. The frontal and parietal lobes had no plaques or tangles, and the frontal cortex exhibited some mild spongiform change. The basal ganglia and thalamus revealed inclusions within the oligodendroglial cells that stained positively to Gallyas silver stain and α-synuclein, and surrounded the nucleus in a cup-like shape.

The cerebellum showed similar inclusions that also were revealed on ubiquitin staining. In addition, occasional granule cells exhibited ubiquitin and α-synuclein positive inclusions. There was loss of Purkinje cells with replacement Bergmann astrocytosis and white matter gliosis. The substantia nigra also showed some vacuolation, loss of pigmented neurones, and some free pigment. No Lewy bodies were seen. Again, there were large numbers of Gallyas and α-synuclein staining inclusions, with occasional whorled-like intraneuronal inclusions visible on α-synuclein staining. The pons was small, especially in the region on the basis pontis, with many oligodendroglial inclusions and some neuronal α-synuclein positive globular inclusions. Sections from the medulla revealed almost complete loss of neurons in the inferior olivary nucleus, with oligodendroglial inclusions again seen on Gallyas, α-synuclein and ubiquitin staining. Occasional neurons also contained α-synuclein positive inclusions. The spinal cord exhibited oligodendroglial inclusions on the Gallyas silver stain.

In summary, the presence of widespread α-synuclein positive oligodendroglial inclusions was diagnostic for multiple system atrophy with no evidence of coexisting Alzheimer's disease or other cerebral pathology.

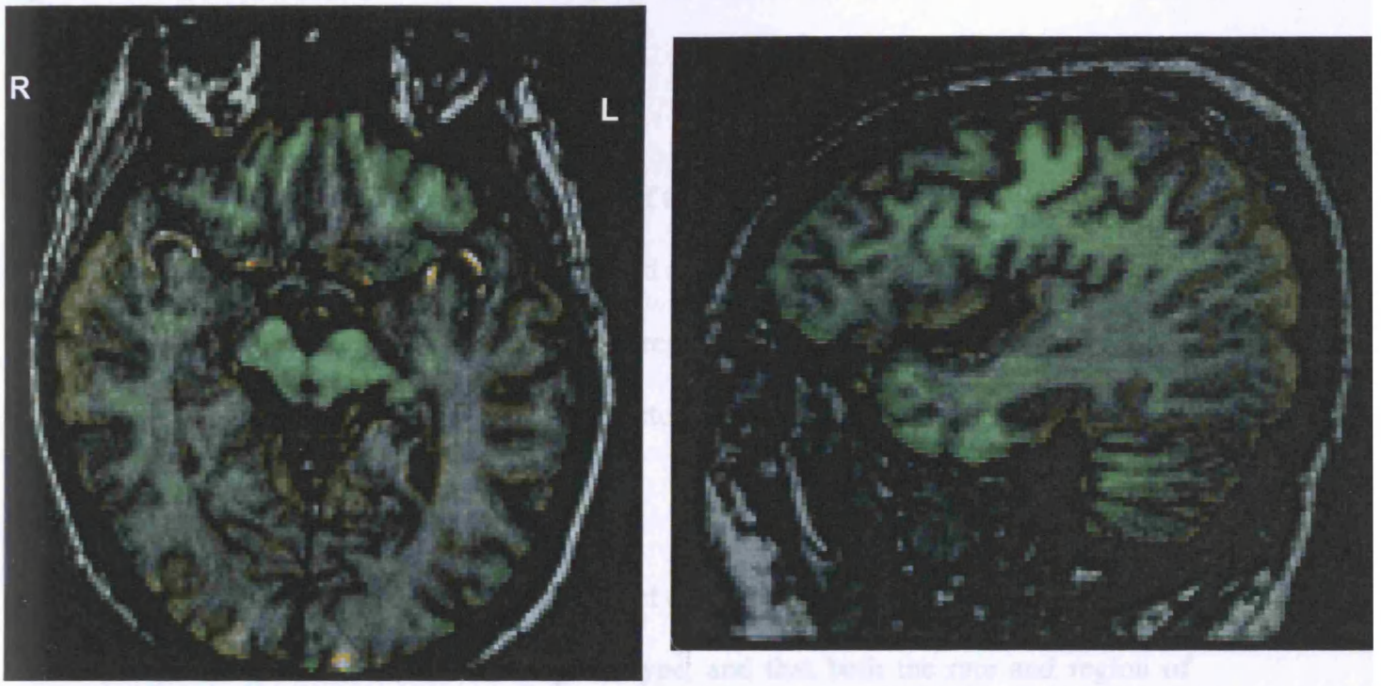
3.5.3.2. MRI results

In the PSP patient, the fluid registered images demonstrated regional volume loss and picked out particularly high rates of atrophy in the midbrain and cerebral peduncles as well as the precentral gyrus compared to the anterior frontal cortex (figure 3.5).
The motor cortex was pathologically more severely affected than the anterior frontal cortex and this is reflected in the volume changes seen on serial imaging, which

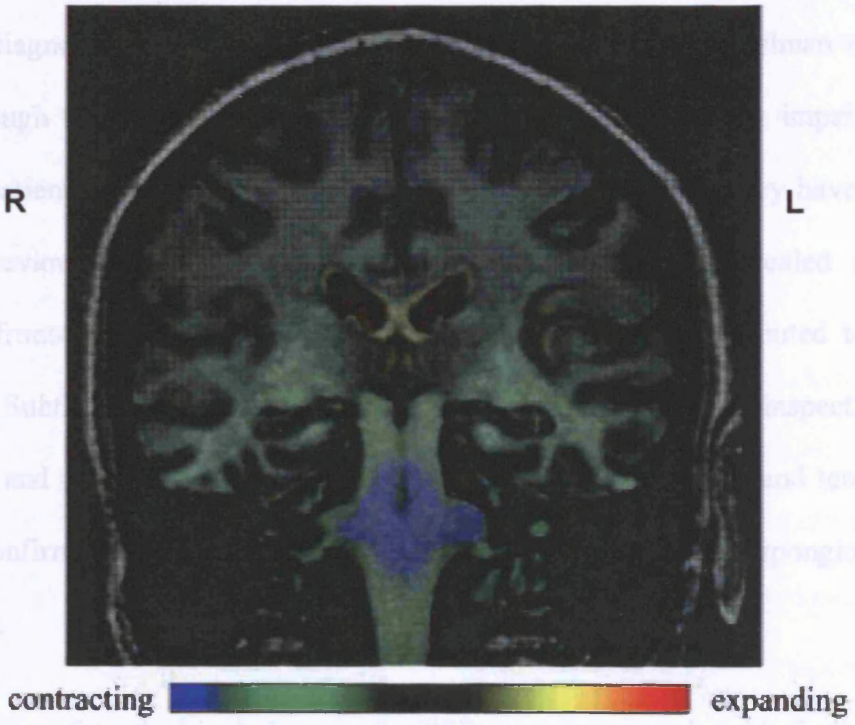
demonstrate a greater atrophy rate in the posterior frontal cortex than the anterior frontal cortex (figure3.5).

In the MSA case, fluid registration of the serial scans revealed striking differences between areas of very rapidly progressing atrophy (e.g., pons and cerebellum) and other areas ostensibly unaffected by atrophy (e.g., parietal and occipital cortices; Fig.3.5). The most prominent atrophy involved the pons and the middle cerebellar peduncles. Increased rates of atrophy were seen also in the midbrain and medulla, vermis and cerebellar white matter, superior and inferior cerebellar peduncles, tegmentum, and olives. There was mild progressive atrophy of the pallidum, but not of the amygdalae, hippocampi, or caudate. The thalami were relatively spared. There was moderate progressive frontal and anterior temporal lobe atrophy, with a notable antero-posterior and infero-superior gradient, extending to involve the genu of the corpus callosum.

Fig 3.5 *Fluid registered imaging in PSP and MSA (pathological correlation in PSP)*



PSP - Axial magnetic resonance imaging scan with voxel-compression mapping overlay mapping changes in the 18 months between scans, demonstrates greatest rates of volume loss in the midbrain. Sagittal magnetic resonance imaging scan with voxel-compression mapping overlay of the right hemisphere, showing volume loss affecting the frontal lobes, particularly the precentral gyrus. Concomitant expansion is seen in the cerebrospinal fluid spaces.



MSA - Coronal MRI scan with voxel-compression-mapping overlay to demonstrate areas undergoing atrophy in the 14 months between scans. Greatest rates of atrophy are demonstrated in the pons and middle cerebellar peduncles and the immediately adjacent midbrain and medulla. Increased atrophy, but at a slower rate, is seen in the upper midbrain and lower medulla. Even slower, but definitely pathological atrophy rates are seen in both temporal lobes. Ventricular enlargement is also shown.

3.5.3.3. Conclusions

Significant and focal progressive atrophy of the same regions most heavily affected at post mortem, were demonstrated using fluid registration of serial MRI scans acquired during life. In the PSP patient, fluid registered serial MRI identified regions undergoing rapid atrophy that closely reflected the anatomical location of cell loss and tau positive NFT load.

These preliminary findings support the tenet of this thesis that the topography of brain tissue loss determines the clinical phenotype, and that both the rate and region of atrophy may be reliably identified using MRI.

Severe cognitive impairment is not a prominent feature of MSA, and dementia excludes a diagnosis of MSA according to the consensus criteria (Gilman *et al.*, 1999). Although the generalised intellectual deterioration and memory impairment seen in the patient described in this chapter are not typical of MSA, they have been described previously (Robbins *et al.*, 1994). Fluid registration revealed subtle progressive frontal and some temporal atrophy that may have contributed to this dysfunction. Subtle atrophy in these cortical areas was seen on gross inspection at postmortem, and pathological involvement of both the anterior frontal and temporal lobes was confirmed by the histological findings of slight superficial spongiosis in these regions.

The distribution of cortical pathology in the PSP case corresponds with findings in previous studies (Daniel *et al.*, 1995; Verny *et al.*, 1996a) which suggested that atrophy is greatest in the posterior frontal lobes. This supports the notion that cortical

atrophy observed on in-vivo imaging is due to primary pathology rather than deafferentation secondary to loss of projecting neurons.

Despite a 19-month interval between the final MRI scan and death in the PSP case, a close correlation was seen between volume loss in-vivo and pathology at autopsy. This suggests that either frontal lobe atrophy occurs relatively early (2-3 years from symptom onset) in the course of the disease (Cordato *et al.,* 2002) and that the pattern of volume loss remains consistent into the latter stages. Serial imaging in the early stages may show accelerated rates of atrophy in areas specific to PSP and MSA, and could prove useful in differential diagnosis of the bradykinetic-rigid syndromes.

We have demonstrated that the focal atrophy pattern revealed using fluid registration seemed to predict that found at postmortem. Additional prospective imaging studies were next carried out to assess the diagnostic value of MR regional volumes and atrophy patterns in the diagnosis of PSP and MSA-P, and to confirm whether atrophy rates are feasible markers of disease progression.

3.5.4. Imaging protocol

One major consideration was the ability of patients to tolerate MRI acquisition. This is dependent upon the time spent in the scanner during which minimal movement is paramount. Scans from patients with a neurodegenerative movement disorder are likely to be subject to movement artefact, limiting their validity. A compromise needs to be reached between better anatomical detail and grey/white matter differentiation being achieved with a longer scan time and the capacity of a patient to tolerate the scan.

Criteria for the scanning protocol included: minimum acquisition time to maximize patient compliance; high signal to noise ratio (SNR); minimum distortion, shading

and susceptibility artifact; high contrast between grey matter and cerebrospinal fluid (CSF) and between grey matter and white matter and; anatomical accuracy (hence resolution). The scan interval also had to be taken into consideration; if too short, it would lead to difficulties detecting the desired signal (atrophy) and if too long would lead to increased attrition due to disease progression and death. A scan interval of around one year was decided upon.

In addition to T1 volumetric imaging, T2 and proton density weighted MRI scans were acquired. These help to identify alternative pathologies such as vascular or inflammatory CNS disease that would exclude a clinical diagnosis of PSP, and exclude healthy controls from the study.

Diffusion weighted MRI scans (chapter 2.4.4 and chapter 8) can be acquired over a very short period of time (less than one minute) and provide information regarding the integrity of normal tissue architecture. Such a short scan acquisition time has considerable advantages in movement disorders.

All MRI scans were acquired on the same 1.5 Tesla (T) GE Signa Unit (General Electric, Milwaukee, WI).

3.5.4.1.T1 weighted imaging

T1-weighted volumetric MR scans were acquired using a spoiled gradient-echo technique with: a 24 x 18cm field of view (FOV); 256mm x 256mm matrix; inversion time (TI)- 650ms; repetition time (TR)-15ms; echo time (TE)-5.4 ms; and flip angle-15°. This yielded 124 contiguous 1.5mm thick slices.

3.5.4.2.T2 and PD weighted imaging

Dual echo proton density and T2 images were acquired using the following parameters. Axial 5mm slices to cover the brain: FOV-250x180mm, 256mm x 256mm matrix, TR-2500ms, TE-30/90ms and, number of excitations (NEX)-1).

3.5.4.3.Diffusion weighted imaging

Axial DWIs were obtained using spin-echo planar sequences (TR- 1000ms, echo time, TE- 99ms, matrix size-128 x 128 pixels, field of view, FOV: 240mm x 240mm, slice thickness 7mm and slice gap 2mm). Diffusion-sensitising gradients were applied perpendicular to the axial slice plane with a b-value of 1000 s/mm^2.

3.6. Recruiting subjects

Recruitment of subjects for the studies incorporated within this thesis was undertaken at the National Hospital for Neurology and Neurosurgery, Queen Square, London. Recruitment was largely from specialist movement disorder clinics (Prof. A.J Lees) but subjects were also recruited from a cognitive disorders clinic (Prof. M.N. Rossor). Prior to inclusion, patients had undergone full clinical assessments. Appropriate investigations such as neuropsychological evaluation, routine haematological and biochemical blood tests and neuroimaging had been performed in order to support the clinical diagnosis. All patients included (and their next of kin where appropriate) were aware of their clinical diagnosis.

Initially the aim was to recruit 25 PSP subjects with 12 MSA-P subjects and 12 PD subjects. As differences in regional volumes and rates of atrophy between the PSP, MSA-P and PD subjects and correlations with the clinical phenotype in these cases was the main interest of this study. MRI scans from healthy control subjects using the same T1 acquisition parameters, acquired for a separate study ongoing within the Dementia Research Centre (UCL) at the National Hospital for Neurology and Neurosurgery, were also used for comparison of disease and healthy control rates of brain atrophy.

3.6.1. Matching of cohorts for age, sex, disease duration or severity

Patients were recruited from outpatient clinics at the National Hospital for Neurology and Neurosurgery. All were required to meet the diagnostic criteria for PSP, MSA and PD. Potential subjects were then informed about the nature of the study and given a patient information sheet outlining its nature as well as the answers to common questions (appendix).

Recruitment was undertaken over a period of several months. Consideration was given to the ability of patients to tolerate the MRI which required that the subject lie flat and still for approximately 20 minutes. Those whose disease had already resulted in severe disability were not included if it was felt that they would be unlikely to tolerate the initial or the follow up scan.

As far as possible, recruitment was directed towards age and sex matching of the groups, and for the PSP and MSA-P subjects, towards matching of disease severity. Parkinson's disease subjects with severe peak dose dyskinesias were not recruited because of concerns about movement artefact.

No financial incentive to take part was offered, but reimbursement of travel expenses was built into the costing of the study.

3.7. Brain donation: addressing the issue early

The accuracy of the clinical diagnostic criteria for these diseases has been reported. Hughes *et al* have reported the accuracy of diagnosis of PSP and other parkinsonian syndromes in the specialist movement disorder service from which the subjects in this study were recruited (Hughes *et al.,* 2002). In their series of 143 cases seen at the National Hospital for Neurology and Neurosurgery, at Queen Square, with Parkinsonism, 19 had a diagnosis of PSP. The positive predictive value, sensitivity, and specificity for this diagnosis were 80.0%, 84.2%, and 96.8%, respectively. They

did not apply specific diagnostic criteria but relied on pattern recognition by movement disorder specialists. The sensitivity and specificity of the diagnostic criteria for PSP in table 3.6a above have been reported at between 50-60% (sensitivity) and 100% (specificity) for probable PSP, and 70-80% (sensitivity) and 90-100% (specificity) for possible PSP. The accuracy of the diagnostic criteria for MSA may not be so high and none have been validated in the context of post mortem verification of the disease (Litvan *et al.,* 2003). An experienced clinicians judgement has been reported as being a better diagnostic marker of the disease, perhaps because it takes into account the results of ancillary investigations. In a large study of post mortem confirmed cases of PD (Hughes *et al.,* 2002), the positive predictive value of the final clinical diagnosis of PD was extremely high, at 98.6%.

Confirmation of the diagnosis, in the absence of a biomarker, relies on post mortem histopathological analysis. The issue of brain donation was broached with all participating subjects.

It was felt that discussion with the subject and their next of kin, at the time of a second consultation, rather than late in the course of the illness was more appropriate for this sensitive topic. All subjects who expressed an interest were given information on the procedure for brain donation and allowed time to formulate and ask questions before consenting. An example of the current information sheet and consent form can be seen in the appendix.

At the time of writing, 11 subjects in total (7 with PSP, 3 with MSA-P and 1 with PD) have consented to brain donation.

3.8. Conclusions

A summary of the proposed study design and methods is outlined in Table 3.8. This design is based upon conclusions drawn from the preliminary study of an extensive

series of brain bank cases of pathologically proven PSP and from the case studies of serial imaging reported in this chapter.

Table 3.8 *Summary of study design and methods overview.*

Recruitment according to clinical diagnostic criteria			✓
		Visit 1	**Visit 2**
History		✓	-
Full clinical examination		✓	✓
Review of inclusion and exclusion criteria		-	✓
Proposed interval (weeks)		-	52
MRI scan:	T1	✓	✓
	T2/Proton density	✓	✓
	DWI	✓	✓
Neuropsychological testing		✓	✓
Bedside cognitive tests	MMSE	✓	✓
	FAB	✓	✓
Clinical rating scales	UPDRS	✓	✓
	HY	✓	✓
	Schwab and England	✓	✓
	Gaze Palsy	✓	✓
Discussion regarding brain donation		✓	✓

Whilst the proposed scan interval was set at 52 weeks, the MRI scanner designated for use in this study was upgraded during the course of the project, meaning that some subjects had to have their proposed second scan date brought forward. This meant a greater variability in scan interval than had been envisaged. In addition, follow up appointments in patients had to be adjusted to ensure that the mean interval was not significantly different between the groups.

Other problems encountered during the project regarding the methods and design of the study, included co-ordinating the neuropsychology appointments and scan appointments. These could not always be undertaken on the same day and so some patients had to return after their scan in order to complete the psychology assessments.

4. The clinical assessment

4.1. Introduction

For inclusion, patients had to have a clinical diagnosis of PSP (possible or probable), MSA-P or PD made after appropriate exclusion of alternatives and based upon the diagnostic criteria described in chapter 3. The patients were also required to be able to tolerate an MRI scan, with no contraindications. Healthy control subject MRI scans were acquired from partners of patients where consent was obtained. The inclusion criteria for controls were that there was no history of neurological disease, cognitive decline, head injury, stroke or psychiatric illness. The number of healthy controls included for comparison of MRI features, was increased by including serial volumetric MRI scans that had been acquired for a separate research study on the same scanner using the same T1 weighted MRI protocol.

4.2. Methods

At the initial visit, consent for participation in the study was obtained. A semi-structured medical history was taken, from the patient and a close informant, focussing on important clinical features of PSP, (determined in part from the study outlined in chapter 3.2.1.) MSA-P and PD.

The semi-structured medical history included details on occupation, current age and handedness, age at symptom onset, estimated disease duration and the presence of: falls (and whether these were backwards); rigidity; tremor (and whether asymmetric or not); bradykinesia; visual symptoms (blurred vision, watery vision, reading difficulties, photophobia, visual hallucinations, eyelid dysfunction); speech and swallowing difficulties; the presence of perceived cognitive dysfunction; postural

dizziness and syncope; bladder dysfunction; drug therapy and levodopa responsiveness.

On clinical examination, the presence of features such as myoclonus, apraxia, dystonia, dyskinesia, cortical sensory loss, alien limb phenomena, pyramidal tract signs and cerebellar signs were recorded.

This information was collected in order to support the suspected clinical diagnosis at the initial visit and to ensure that at the follow up visit, no features excluding a diagnosis of PSP, MSA-P or PD were present. Any revision of the clinical diagnosis was recorded if appropriate the patient was excluded from further data analysis.

A single rater (myself) undertook all clinical assessments.

4.2.1. Clinical rating scales

All subjects had the MMSE, see Appendix (Folstein *et al.,* 1975) recorded in addition to scores on other clinical rating scales detailed below.

4.2.1.1. UPDRS

The Unified Parkinson's Disease Rating Scale (UPDRS) (Fahn S, 1987) as described in section 1.2.8.1 was designed to provide a semi-quantitative measure of the signs and symptoms of Parkinson's disease (PD) in clinical practice and research. Because clinicians are familiar with the scale and because the motor examination (part III) of the UPDRS has been validated as a method of assessing function in PSP (Cubo *et al.,* 2000) and in MSA (Tison *et al.*, 2002), this scale was chosen as a method of assessing various aspects of disease severity. To recap, part one of the scale assesses mentation behaviour and mood, part two, activities of daily living (ADL) and part four, complications of therapy. A more detailed assessment of mentation behaviour and mood was obtained from the more formal neuropsychological testing (chapter 5). The

score from part four of the UPDRS in PSP and MSA is difficult to compare with PD as there is little if any response to levo-dopa in most MSA and PSP cases and dyskinesias are much less common. For these reasons, comparison between the three disease groups was limited to scores on the UPDRS II and III.

4.2.1.2. Hoehn and Yahr

The Hoehn and Yahr staging of disease has five levels of severity (Hoehn and Yahr, 1967): stage 1- unilateral disease; stage-2 bilateral disease without impairment of balance; stage 3- mild to moderate bilateral disease, some postural instability, physically independent; stage 4- severe disability, still able to walk or stand unassisted; and stage 5- wheelchair bound or bedridden unless aided (appendix, chapter 14.5). This has the advantage of being a simple and easily applicable method of grading disease severity.

4.2.1.3. Schwab and England ADL score

The Schwab and England ADL score, measures disease severity in terms of activities of daily living (ADL). A score of 100% means complete independence and 0% is a bedridden state with dependency for all activities including swallowing, bladder and bowel functions (appendix). Allocation to a particular percentage level is dependent on subjective interpretation of an individuals level of disability and very few will fit perfectly with the description at each percentage level. Nevertheless, an assessment of functioning based on this score is widely accepted as a measure of disability and many physicians are familiar with its use.

4.2.1.4. Frontal assessment battery

The frontal assessment battery (FAB) is a bedside test, focussing on frontal/sub-cortical cognitive impairment (Dubois *et al.,* 2000). It can be administered in less than ten minutes and consists of six sections, each with a maximum score of 3 points. It

has been validated as a bedside test of frontal cognitive function and shown to correlate with more formal neuropsychological tests. A detailed analysis of the FAB and its relation to formal neuropsychological test scores in the patients recruited to this thesis can be found in chapter 5.

4.2.1.5. Gaze palsy score

A supranuclear vertical gaze palsy is one of the cardinal features of PSP, and yet as described in section 1.2.8.1, there is no validated bedside scale. The severity of the gaze palsy varies from individual to individual and its onset during the course of the disease is variable. The neurological control of eye-movements is complex. Many brainstem, basal ganglia and cortical areas have roles in the generation of normal eye-movements. Abnormalities of saccadic eye movements occur early in PSP (Leigh and Riley, 2000) and this is reflected in the diagnostic criteria. In contrast, abnormalities of saccadic eye movements in MSA-P and PD are less common, less severe and occur later in the course of the illness. Many different anatomical structures or pathways are involved in the generation of normal saccadic eye movements (for a review see (Leigh and Kennard, 2004).

In PSP, pathological involvement of brainstem nuclei including those in the rostral interstitial medial longitudinal fasciculus (riMLF) contributes to slow vertical saccades and pathology in the para-pontine reticular formation to slow horizontal saccades. A schematic diagram of the brainstem saccadic generators is shown in figure 4.1. Accurate measurement of saccadic eye-movements is best undertaken using electro-oculography or by using infrared video recording of eye movements. These recording methods are either moderately invasive or rely on expensive equipment and neither was undertaken in this thesis.

However, accurate clinical recording of the severity of any gaze palsy was felt to be important in the PSP patients. For this reason, and with a view to relating the gaze palsy severity to the imaging features of the disease, a novel scale (devised specifically for a large multi-centre clinical trial of Riluzole in PSP and MSA, the NNIPPS trial) based on the clinical assessment of gaze palsy was used (appendix). Grades of normal, mild, moderate or severely hypometric (score 0, 1, 2, 3) or slow (score 0, 1, 2, 3) saccades in the vertical (up), vertical (down) or horizontal planes were added to scores of no, mild, moderate or severe eyelid dysfunction (0, 1, 2 or 3) to give a maximum severity score of 21. This is a pragmatic and quickly administered scoring system. As yet, it has not been validated by comparison with accepted methods of assessment such as electro-oculography, nor has inter-rater reliability been assessed. The scale was incorporated specifically to detect the more severe vertical gaze palsy typical of PSP. Abnormalities associated with MSA such as square wave jerks and nystagmus, do not feature and would not have affected the gaze palsy score in this group of patients.

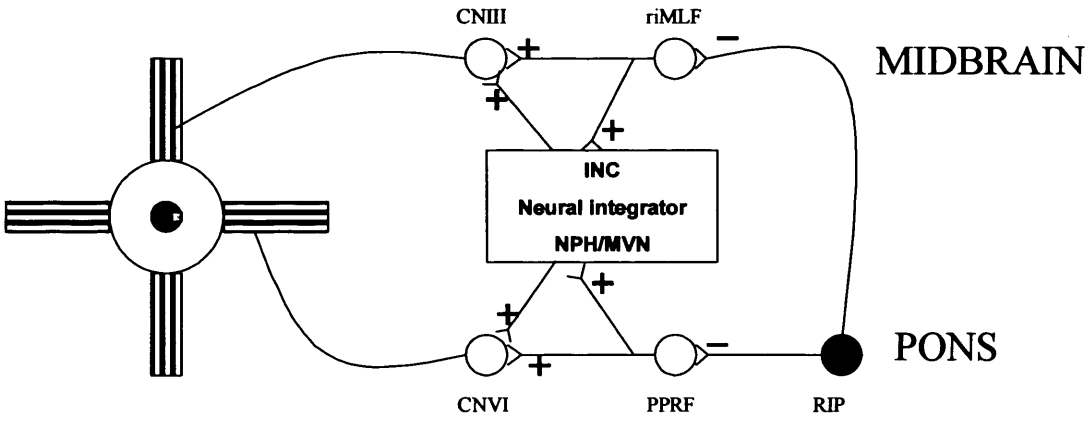


Figure 4.1 *Schematic summary of brainstem saccadic generators.*
Horizontally acting extra-ocular muscles, such as the lateral rectus, are innervated by motor neurons in the abducens nucleus (CN VI); these motor neurons receive their saccadic commands from the PPRF. Vertically acting extra-ocular muscles, such as the superior rectus, are innervated by motor neurons in the occulomotor nucleus (CN III) and trochlear nucleus (not shown); these motor neurons receive saccadic commands from burst neurons in the riMLF in the midbrain. Burst neurons in both the PPRF and riMLF receive inhibitory inputs from omnipause neurons, which lie in nucleus raphe interpositus (RIP) in the pons. When omnipause neurons cease discharge, burst neurons in the PPRF and riMLF generate a saccadic pulse, which is passed to the motor neurons and causes a phasic contraction of the extra-ocular muscles to move the eye rapidly in a saccade. Horizontal saccadic pulse signals are integrated by a network consisting of neurons in the medullary nucleus prepositus hypoglossi and adjacent medial vestibular nucleus (NPH/MVN); vertical saccadic pulses are integrated in the midbrain interstitial nucleus of Cajal (INC) for vertical saccades. The output of this neural integrator is a step signal that makes it possible for extra-ocular muscles to sustain a tonic contraction and hold the eye in its new position (Leigh and Kennard, 2004).

4.2.2. General clinical examination

In addition to the structured assessments of disease severity, a full neurological examination was carried out at the initial assessment.

4.2.3. Statistical analysis

Tests of skewness and kurtosis were performed to assess the normality of the data. For normally distributed data, a one-way analysis of variance was performed and where significant differences were detected ($p<0.05$), post hoc analysis was undertaken to determine between which groups (PSP, MSA-P, PD or healthy controls) the differences occurred. For non-normally distributed data, a Mann-Whitney U test was performed to assess differences between group mean values. A Chi-squared test was used to assess sex distribution between groups. The relationship between these measures of disease severity and disease duration, at the time of the first assessment, in PSP, MSA-P and PD was assessed using a bootstrapped linear regression analysis with age as a co-variable.

4.3. Results

Twenty-four patients with PSP, eleven with MSA-P, twelve with PD and eighteen healthy controls were included. Three of the PSP subjects were then excluded from further follow up as the first scan was of poor quality because of excessive movement artefact. These patients were not included in any of the data analyses. One PSP and one MSA-P patient failed to attend for a second scan a year later because of disease progression. Two PD patients withdrew from the study after the initial scan but gave consent for their data up to the point of withdrawal to be included. One patient with MSA-P and one with PSP died before the follow up appointment. One patient classified as PSP at the initial visit on the basis of the diagnostic criteria developed

visual hallucinations and a clinical syndrome more in keeping with a clinical diagnosis of dementia with Lewy bodies (DLB) and was excluded from all data analysis. Follow up scans in one patient with PSP and one with PD were of poor quality and were not included in the data analysis. Of the 65 subjects recruited, 21 patients with PSP, 11 with MSA, 12 with PD and 18 healthy controls were included for analysis of clinical data from the first visit (total 62). 18 PSP, 9 MSA, 10 PD and 18 healthy controls (total 55) were included for analysis of follow up clinical data, and 17 with PSP, 9 with MSA, 9 with PD and 18 healthy controls, for analysis of follow up imaging data. These data are reported in table 4.1. The numbers of subjects included in imaging analysis and neuropsychological test score analysis are reported later.

There were no age or sex differences between the groups, disease duration was significantly longer in the PD group ($p<0.001$) and motor deficit was less severe in the PD group according to UPDRS II, III and H+Y scores. There was no difference between the PSP and MSA-P groups. At the time of the initial visit, the MMSE was lower in the PSP, MSA-P and PD groups than in healthy controls, with no significant differences between the diseased patients. The FAB score discriminated PSP from MSA-P and from PD (this is expanded on in chapter 5).

The gaze palsy score was markedly increased in the PSP patients. The group differences at the initial visit were consistent with those found at the time of follow up.

Bootstrapped regression analysis did not reveal any significant relationship between disease duration and any markers of disease severity.

Table 4.1. *Clinical features (Mean scores and SD)*					
	PSP	**MSA**	**PD**	**Control**	**P value**
n	21	11	12	18	
Sex (M:F)	14:7	6:5	8:4	10:8	0.8
Age (yrs)	65.4 (6.2)	62.0 (7.6)	65.5 (9.2)	66.8 (5.4)	0.3
Disease duration (yrs)	4.5 (1.5)	5.4 (1.6)	13.3 (6.7)	-	PD vs. PSP/MSA, p<0.001
First visit					
UPDRS2	20.2 (6.9)	24.9 (7.0)	13.4 (4.6)#	-	<0.001
UPDRS3	21.0 (7.9)	26.8 (9.7)	16.7 (5.1)*	-	0.01
HY	3.5 (0.6)	3.9 (0.8)	2.7 (0.5)♦	-	0.001
MMSE	26.0 (3.1)	26.4 (3.2)	27.7 (2.5)	29.5 (0.6)	C vs. PSP/MSA, p<0.001, C vs. PD, p=0.009
FAB	11.7 (3.6)	14.7 (2.5)	16.5 (1.4)	-	PD vs. PSP, p<0.001 PSP vs. MSA, p=0.02
Schwab and Eng.	59.5 (23.9)	53.4 (21.6)	78.3 (12.7)	-	PSP vs. PD, p=0.02 MSA vs. PD, p=0.005
Gaze palsy score	12.2 (5.9)	2.4 (2.3)	0 (0)	-	PSP vs. MSA/PD, p<0.001 MSA vs. PD, p<0.001
Interval (days)	292 (59)	262 (66)	268 (52)	252 (40)	PSP vs. C, p=0.02
Second visit					
n	18	9	10	18	
UPDRS2	28.2 (7.8)	30.4 (7.6)	17.9 (5.7)$	-	<0.001
UPDRS3	29.0 (7.9)	32.0 (10.3)	20.6 (7.2)▾	-	0.01
HY	4.1 (0.7)	4.2 (0.8)	3.1 (0.6)♠	-	0.001
MMSE	26.2 (3.1)	26.4 (3.4)	27.7 (2.2)	-	0.4
FAB	10.3 (4.2)	13.2 (3.3)	15.8 (2.4)†	-	0.002
Schwab and Eng.	47.2 (21.1)	41.1 (21.5)	77.0 (8.2)	-	PSP vs. PD, p<0.001 MSA vs. PD, p=0.001
Gaze palsy score	13.9 (5.8)	4.1 (4.1)	0.1 (0.3)	-	PSP vs. MSA/PD, p<0.001 MSA vs. PD, p<0.001
Reported p values are from ANOVA or from Mann-Whitney U test. Post hoc (bonferroni) tests- #PD vs. MSA, p=0.001, PD vs. PSP, p=0.03; *PD vs. MSA, p=0.01; ♦PD vs. PSP, p=0.01, PD vs. MSA, p=0.001; $PD vs. MSA, p=0.002, PD vs. PSP, p=0.003; ▾PD vs. MSA, p=0.02, PD vs. PSP, p=0.05; ♠PD vs. PSP, p=0.002, PD vs. MSA, p=0.004; †PD vs. PSP, p=0.002.					
UPDRS-unified Parkinson's disease rating scale; HY- Hoehn and Yahr; MMSE-mini mental state examination; FAB- frontal assessment battery.					

4.4. Conclusions

The groups are not significantly different in terms of age and sex distribution. There are however significant differences in scores on the clinical rating scales, and bedside tests of cognitive function between the patients with PSP and MSA-P and those with PD.

Disease duration is significantly longer in the PD patients. This is because an attempt was made to recruit subjects with more severe disease, in order to achieve closer matching with the PSP and MSA-P subjects. There is no significant difference in the

interval between assessments in the PSP, MSA-P or PD patients. The interval between assessments is longer in the PSP group than the healthy control group. However progression in clinical scores and imaging findings was standardised to an annual interval or the interval used as a covariate in all analysis.

That the mean UPDRS II and III scores were significantly different between PSP and PD, and MSA-P and PD patients is not surprising. The UPDRS II assesses activities of daily living (ADL). Speech, swallowing, hygiene and falling scores are likely to be higher in PSP and MSA-P than in PD. On the UPDRS III, speech, finger taps, arising from a chair, gait and postural stability scores are likely to be greater in PSP and MSA-P than in PD. These factors account for the greater disease severity scores in PSP and MSA-P. The Hoehn and Yahr is an assessment of motor disability and if there is postural instability, a score of at least 3 must be recorded. This means that all PSP and most of the MSA-P patients will have a score of at least 3. PD patients on the other hand are less likely to have impaired balance and so may score lower.

The gaze palsy score differentiates the groups well, as would be expected. No saccadic eye movement abnormalities were detected in the PD patients.

The FAB also differentiates the groups suggesting it can help to detect the more severe sub-cortical cognitive impairment in PSP. This is discussed further in chapter 5.

There were no clear associations detected between disease duration and measures of disease severity. This is surprising as it is clear that on an individual level, these diseases result in greater disability as they progress. This lack of association could be explained, firstly by the heterogeneity of the diseases studied, resulting in considerable variation in individual disease duration from onset to increasing disability and death in PSP and MSA-P. Secondly, the start of the disease process was

recorded at the time of the first symptom, taken from the history. This is subjective and, in some cases likely to be an inaccurate measure of onset, which will affect any association between disease duration and severity. The clinical rating scales used have inevitable ceiling effects. Hence once a patient with PSP is falling more than once per day, or is unable to arise from a seated position without help, then they cannot increase their score for that particular aspect of the UPDRS assessment regardless of true disease progression.

Despite efforts to match the patient groups, disease severity measured using established and new disease rating scales in PSP and MSA-P patients was greater than in PD patients.

In summary, all patient groups and healthy controls were relatively well matched in terms of age and sex. PSP and MSA-P patients were matched for disease severity; however aspects of cognitive function assessed according to the FAB were more severe in the PSP patients.

5. Neuropsychological assessments

5.1. Introduction

The neuropsychological test battery was selected specifically in order to detect the presence of subcortical cognitive impairment, but also other aspects of more general cognitive function (Chapter 3.5.2).

The objectives behind analysis of the neuropsychological scores were: (1) to assess the subjects with PSP, MSA-P and PD, at entry and at a follow up assessment; (2) to assess the change in neuropsychological test scores over time, in the different disease groups; (3) to assess the ability of the FAB to quantify executive dysfunction in bradykinetic-rigid disorders as well as; (4) determining the correlation between the FAB and more extensive tests of executive function and cognitive function in order to evaluate the practical value of the FAB in differentiating these bradykinetic-rigid syndromes.

A total of 40 individuals completed or partially completed the neuropsychological test battery at the time of initial assessment. These included 17 with PSP, 11 with MSA-P and 12 with PD. A total of 38 subjects attended for a second neuropsychological assessment, including 18 with PSP (3 with no initial assessment), 9 with MSA-P and 10 with PD. In total, 33 subjects completed neuropsychological assessments at entry and at follow up, 15 with PSP, 9 with MSA-P, and 9 with PD.

5.2. Neuropsychology profiles at initial assessment.

5.2.1. Methods

Neuropsychological testing was completed at the time of the clinical assessment and undertaken by a trained neuropsychologist. The clinical tests were completed with no knowledge of the results of the formal neuropsychological evaluation.

Patient characteristics were compared using a one-way analysis of variance for parametric data and for non-parametric data, Chi-squared or Mann Whitney U tests were used.

Neuropsychology test scores at the first visit were standardised:

z-score = (x-mean)/SD

Where, x is the raw neuropsychology score, mean is the mean overall score for the entire cohort and SD is the standard deviation of the group.

Group differences were determined using one-way analysis of variance (ANOVA) with post-hoc analysis for parametric data and a Mann-Whitney U test for non-parametric data.

5.2.2. Results

The groups did not differ significantly in terms of age or gender distribution ($p>0.05$). Duration of illness, severity of motor symptoms (UPDRS III) and disability in activities of daily living (UPDRS II) were significantly different across the groups (table 5.1). Post-hoc analyses showed that the significant group differences were due to the patients with PSP and MSA-P having more severe motor symptoms and greater disability than the PD group; whereas the UPDRS II and III ratings for the MSA-P and PSP patients did not differ significantly. The PD patients had a significantly longer duration of illness than those with PSP or MSA-P. There was no significant difference in disease duration between PSP and MSA-P.

Table 5.1 *Mean (SD) subject details for those attending first psychology assessment*

	n (M:F)	age (yrs)	disease duration (yrs)	UPDRS2	UPDRS3	H+Y
PSP	17 (12:5)	64.8 (5.9)	4.4 (1.6)	20.1 (6.4)	20.6 (7.2)	3.5 (0.7)
MSA-P	11 (6:5)	61.9 (7.6)	5.4 (1.6)	24.9 (7.0)	26.8 (9.7)	3.9 (0.8)
PD	12 (8:4)	65.4 (9.2)	13.3 (6.7)♦	13.9 (5.1)*	16.7 (5.1)**	2.8 (0.6)♠
P Value	0.7	0.5	<0.001	<0.001	0.009	0.002

♦-p<0.001 vs PSP and MSA; *-p=0.04 vs PSP, p<0.001 vs. MSA; **-p=0.008 vs MSA; ♠-p=0.02vs PSP,p=0.001 vs MSA

For normally distributed data, a one way analysis of variance (ANOVA) comparing standardised neuropsychology scores (z-scores), between the PSP, MSA-P and PD subjects detected significant differences in the following: scores on the vocabulary subtest of the WAIS-R were significantly lower in PSP compared to PD subjects ($p=0.03$); the PSP group made more errors than PD subjects on the WCST ($p=0.03$); and scores on the verbal fluency tests in PSP were also lower than in PD ($p=0.001$).

For non-parametric data, the Mann-Whitney U test detected group differences in: the Mattis DRS total score, which was significantly lower in PSP than in MSA-P ($p=0.01$) and in PSP compared to PD ($p=0.03$); the DRS initiation and perseveration sub-test scores which were lower in PSP compared to PD ($p=0.04$); the DRS construction sub-test which was lower in PSP compared to PD ($p=0.007$); the number of categories sorted on the WCST which was lower in PSP compared to MSA-P ($p=0.003$); and in the apathy scores which were greater in PSP compared to PD ($p=0.03$) and in MSA-P compared to PD ($p=0.006$). Differences in other neuropsychological test scores between groups did not reach significance at the 5% level.

5.2.3. Conclusions

Significant differences in the neuropsychological test scores between the patient groups suggest more severe cognitive impairment in PSP. The Mattis DRS subscores which are influenced most by executive function (DRS-initiation/perseveration and DRS construction) were lower in PSP than in PD.

While executive dysfunction is described in some cases of PD (Levin and Katzen, 1995), it is much more severe in PSP (Maher *et al.*, 1985). The results of the

psychological profiles of the patients studied for this thesis are in keeping with these previous observations.

PSP, MSA-P and PD all have pathological dopaminergic depletion in the substantia nigra and other basal ganglia neurons but there are important and widespread neuroanatomical and neurochemical differences between these conditions. Alexander *et al* proposed that three of five cortico-subcortical circuits are affected in PSP (Chapter 3.4.1). They suggested that the preferentially affected circuits were, the dorsolateral prefrontal circuit which is involved in cognition, the lateral orbitofrontal circuit which is involved in social affect, and the anterior cingulate circuit which is involved in attention (Alexander *et al.,* 1986; Alexander and Crutcher, 1990). In PSP, as well as damage to the striato-frontal pathways due to pathology in the basal ganglia, the prefrontal cortex is directly involved (Verny *et al.,* 1996a). This may help to explain some of the differences in the results of formal neuropsychological testing, observed in these degenerative basal ganglia disorders (Testa et al, 1993; Lange et al, 1993; Robbins et al, 1994; Pillon et al, 1995; Meco et al, 1996; Soliveri et al, 2000; Dujardin et al, 2004). In general, these previous studies using a variety of standardized batteries of cognitive tests or the Cambridge automated neuropsychological test battery (CANTAB) have established that the differences in the cognitive profile of patients with idiopathic PD, PSP or MSA are mainly quantitative rather than qualitative (i.e. the test scores differ at a group level but this is difficult to appreciate at the bedside). Across studies, patients with PSP show the most severe and extensive cognitive deficits particularly on tests of executive function and patients with PD the least severe deficits. Nevertheless, some qualitative differences in the profile of cognitive dysfunction in the three Parkinsonian groups have also been reported. For example, simple and alternating semantic and phonemic verbal fluency

tests have been shown to correctly differentiate PSP, PD and MSA and are considered sensitive measures to aid in the differential diagnosis (Lange *et al.*, 2003).

The differences between patient groups observed in this thesis are similar to those reported in previously, demonstrating that PSP patients performed poorly on the Mattis DRS and the WCST compared to MSA and PD patients (Pillon *et al.*, 1995). In contrast, other studies comparing the cognitive profile of PSP with that seen in MSA-P and PD identified clear differences in verbal fluency not replicated in this thesis (Lange *et al.*, 2003).

The number of patients in each disease group in this study compares favourably with those recruited to previous studies and the test battery also bears similarities. However, the cohort was relatively small and the larger number of patients with MSA-P in particular, may have allowed identification of further differences between the disease groups.

5.3. Change in Neuropsychology Profiles

5.3.1. Introduction

PSP, MSA-P and PD are all progressive neurodegenerative diseases. It is expected that the cognitive dysfunction observed progresses more rapidly in PSP than in MSA-P or PD. Few studies report the results of neuropsychological follow up in patients with PSP, compared with MSA-P and PD. The results of those that do are often confounded by the fact that the patients most affected, who progress rapidly, drop out from the study (Soliveri *et al.*, 2000).

The results of neuropsychological testing at follow up were compared with those at baseline in patients with both sets of data.

5.3.2. Methods

The change in psychology test scores at first and second assessments were corrected for differences in interval: corrected difference= difference × (365 ÷ Interval in days). A one-way analysis of variance was used to look for significant group differences in normally distributed data, and for non-parametric data, Mann-Whitney U analysis was used.

5.3.3. Results

In total, 37 subjects completed neuropsychology tests at the time of their second appointment. Two subjects with PD dropped out of the study, and a further patient with PD did not complete the neuropsychology testing. One patient with MSA died and another was too ill to attend. Three subjects with PSP who did not complete neuropsychology testing at the first appointment attended at the time of their second appointment. Of the 17 who attended the first appointment, one PSP patient subsequently died and another was too ill to attend.

Table 5.2 *Mean (SD) subject details for those attending second psychology assessment.*

	n (M:F)	Age (yrs)	Disease duration (yrs)	UPDRSII	UPDRSIII	H+Y
PSP	18 (11:7)	65.4 ± 6.3	4.6 ± 1.6	28.2 ± 7.8	29.0 ± 7.9	4.1 ± 0.7
MSA-P	9 (5:4)	62.3 ± 8.0	5.4 ± 1.7	30.4 ± 7.6	32.0 ± 10.3	4.2 ± 0.8
PD	10 (7:3)	64.4 ± 7.8	12.0 ± 5.0♦	17.9 ± 5.7*	20.6 ± 7.2**	3.1 ± 0.6♠
P value	0.8	0.6	<0.001	0.003	0.01	0.001

♦-p<0.001 vs. PSP and MSA; *-p=0.003 vs. PSP, p=0.002 vs. MSA; **-p=0.02 vs. MSA, p=0.05 vs. PSP; ♠-p=0.002 vs. PSP, p=0.004 vs. MSA

Table 5.3 *Mean (SD) subject details and annualised severity score differences for those attending both psychology assessments.*

	n (M:F)	Age (yrs)	Disease duration (yrs)	UPDRSII change	UPDRSIII change	HY change	MMSE change	FAB change
PSP	15 (10:5)	63.6 (5.0)	6.9 (4.5)	12.0 (7.7)	13.4 (11.3)	0.9 (1.0)	0.8 (3.0)	3.1 (4.6)
MSA-P	9 (5:4)	62.3 (8.0)	5.4 (1.7)	8.9 (6.5)	10.3 (7.7)	0.7 (0.9)	-0.9 (3.1)	1.8 (3.2)
PD	9 (6:3)	64.3 (8.3)	12.9 (4.3)♦	7.3 (4.5)	7.6 (9.8)	0.5 (0.7)	0.7 (2.9)	1.2 (2.1)
P value	0.8	0.8	p<0.001 vs PSP and MSA	0.3	0.3	0.7	0.4	0.6

♦-p<0.001 vs PSP and MSA

Data for patients attending both neuropsychology assessments is shown in table 5.3. The only significant difference in the groups was in the delayed recall subtest of the Rey AVLT (Mann-Whitney U, p=0.006). In this test, the difference in scores between the first and second assessments was greatest in the MSA-P group with a significant deterioration compared to the PD group and a trend towards a difference in comparison to the PSP group (p=0.07). No other significant differences between the rates of change in neuropsychological tests, in the different disease groups were found.

5.3.4. Conclusions

The PSP group showed greater change on all tests, however, no significant differences in the mean change in neuropsychology test scores between groups were noted. Soliveri *et al* detected more rapid decline in tasks related to attention, set shifting and categorisation in PSP, in a similarly sized sample of patients with PSP, MSA-P and PD (Soliveri *et al.,* 2000), although the interval between testing was longer (18-21 months compared to 6-12 months in this study). Another factor is differential drop out. If the subjects who progress fastest drop out differently between groups then a real effect may be lost. Hence, the smaller number of subjects with complete neuropsychological test scores, and the effects of differential drop out, particularly in MSA may have contributed to the lack of a significant difference.

Neuropsychology tests are prone to problems with floor and ceiling effects. If a test score is already at its limit, then no change can be observed. This may be reflected in the PSP group results. Reviewing the individuals recruited to this study on a subjective basis certainly gives the impression that there was a progressive decline in the cognitive deficit as well as the motor disability, and it was surprising that no statistically significant group differences are observed. The change in the psychology

scores is unlikely to be linear, whereas applying a correction for an annualised interval presumes that it is. Whilst this is not an ideal method of normalising the groups, it standardises the change.

The diseases studied are heterogeneous and, progress at varying rates (Chapter 3). They also differ in the degree of cognitive impairment and its evolution. In PSP, this was recognised in the first description of the disease (Steele *et al.,* 1964). This could help to explain why more profound differences on initial assessment or on serial testing were not detected between the disease groups.

Formal neuropsychological tests are useful and have a place in both diagnosis and the clinical management of these conditions. However for the clinician, a reliable and repeatable bedside test of executive function would have utility as an adjunct to the other clinical examination findings, if it were able to help discriminate PSP from other degenerative bradykinetic rigid syndromes. One such test with this potential is the frontal assessment battery (FAB, appendix).

5.4. The Frontal Assessment Battery

5.4.1. Introduction

The frontal assessment battery (FAB, appendix, (Dubois *et al.,* 2000) has been described briefly in chapters 2 and 3. It consists of a sequence of six subtests and has been designed specifically as a quantitative bedside test of sub-cortical/executive dysfunction.

The six subtests (scoring between 0-3 on each) explore: *conceptualisation* (the similarities subtest); *mental flexibility* (verbal fluency for the letter 'S'); *motor programming* (a Luria motor sequence, "fist-edge-palm"); *sensitivity to interference*

(the conflicting instructions subtest); *inhibitory control* (the go-no-go subtest); and *environmental autonomy* (a test of prehension behaviour).

The testing of these features at the bedside allows for simple quantification of a patients ability to conduct appropriate goal directed behaviour and for adapting their response to new or challenging situations. These functions are mediated by the pre-frontal cortex, and it would be expected for patients with PSP to perform poorly compared to those with PD or MSA-P, in whom there is pathological involvement of the basal ganglia, but not extensive involvement of the pre-frontal cortex. The scores on the six subtests add up to give a maximum of 18 (indicating no executive dysfunction). The scale has been well validated in PSP, MSA-P and PD (Dubois *et al.*, 2000), as well as FTD and AD (Slachevsky *et al.*, 2004).

5.4.2. Methods

A single examiner (myself) recorded the FAB scores, prior to the patients formal neuropsychological testing. Mean FAB scores in the different diseases were compared using a one-way analysis of variance with post-hoc t-tests. In order to assess whether difference in the FAB scores between the disease groups was influenced by disease severity or duration, group differences were assessed with disease duration and the UPDRS III score as covariates. Additionally, patients were divided into those with PSP and those with MSA-P or PD and these two groups subdivided into those with a Hoehn and Yahr score >3 or ≤ 3. Mean FAB scores at both levels of disease severity were compared.

In the whole cohort, the cut off score on the FAB that best discriminated PSP from MSA-P or PD was determined by applying a receiver operator characteristic (ROC) curve analysis.

5.4.3. Results

The FAB took approximately 10 minutes to complete in each patient. FAB scores were significantly different across the three groups ($p<0.001$, table 5.4). The mean FAB scores of the PSP patients were significantly lower than those of the MSA or PD patients ($p=0.02$ and $p<0.001$ respectively). FAB scores of the MSA group were only just significantly lower than those of the PD group ($p=0.047$).

Table 5.4 *Number of subjects with FAB scores in the 'normal' and 'abnormal' range and mean (SD) FAB scores in the PSP, MSA-P and PD groups.*

	PSP	MSA-P	PD
Mean (SD) FAB	**11.7 (3.0)*##**	**14.7 (2.5) #**	**16.5 (1.4)**
50th Centile	**15**		
FAB ≥15	**3**	**7**	**11**
FAB <15	**14**	**4**	**1**

*p=0.02 vs. MSA; ##p<0.001 vs. PD; #p=0.047 vs. PD

The group differences in FAB scores remained significant even after controlling for duration of illness and disease severity ($p<0.001$).

Applying a non-parametric test (Mann-Whitney U) of group differences in the FAB for two different levels of physical disease severity based on the Hoehn and Yahr scores, confirmed that FAB scores in PSP patients with Hoehn and Yahr of less than or equal to 3 ($n=10$) compared to those with PD/MSA-P with Hoehn and Yahr at the same level ($n=11$) were significantly different ($p=0.003$). FAB scores for Hoehn and Yahr of 4 or more were also compared between PSP and PD/MSA ($n=7$ and $n=8$) and remained significantly different ($p=0.009$).

A receiver-operating characteristic (ROC) curve analysis showed that a cut-off of less than 15 discriminated PSP from MSA-P and PD with a sensitivity and specificity of 78%. The area under the ROC curve was 0.84 indicating that the FAB had good discriminant ability.

Subsequently, scores of less than 15 were considered to reflect greater executive dysfunction. The proportion of patients with FAB scores below 15 or scores ≥ 15

were compared across the three groups (table 5.4). While in the PSP group 14 patients (82%) had FAB scores in the "low" range, only 4 MSA-P (36%) and 1 PD (8%) patients had scores below 15. The differences across the groups were significantly different (Chi-Squared =16.12, df=2, $p<0.001$).

Of the six FAB sub-scores, differences between the PSP, MSA-P and PD groups were significant for the lexical fluency ($p=0.001$), motor series ($p=0.001$), conflicting instructions ($p=0.03$), and prehension behaviour ($p=0.02$), but not the similarities or go no go subtests ($p=0.08$ and $p=0.4$ respectively). Post-hoc analyses showed that the PSP and MSA-P groups differed significantly on the lexical fluency and motor series ($p=0.002$ for both), whereas the PSP and PD groups differed on similarities ($p=0.04$), lexical fluency ($p=0.002$), motor series ($p=0.001$), conflicting instructions ($p=0.01$) and prehension behaviour ($p=0.009$) but not the go-no-go subtest ($p=0.2$).

A stepwise discriminant function analysis using the six FAB subtests showed that the lexical fluency and motor series sub-scores correctly classified 70% of the PSP, MSA and PD patients (motor series: Wilk's lambda =0.611, $F[2,37]=11.7$, $p<0.001$; lexical fluency: Wilk's lambda =0.50, $F[2,37]=7.5$, $p<0.001$). Predicted group membership on the basis of these two FAB sub-scores was accurate for 76.5% of the PSP, 75% of the PD and 54.5% of the MSA patients.

5.4.4. Conclusions

The FAB scores differentiate PSP, MSA-P and PD, with the PSP group having the lowest scores and the PD patients the highest. Discriminant function analysis confirms that 70% of the patients were correctly assigned their clinical diagnosis of PSP, MSA-P or PD on the basis of their FAB scores alone. Furthermore, while 82% of the PSP patients had FAB scores of less than 15, such scores were only found in 36% of the MSA and 8% of the PD group. In a stepwise multiple regression analysis, alternating

semantic fluency was found to be the variable that accounted for 80% of the variance of the FAB scores.

In the original description of the FAB, no difference between the score in MSA and PD was detected (Dubois *et al.,* 2000). This may have been because of the limited number of MSA patients recruited (n=6). The PSP patients (n=47) in the Dubois study had a mean (SD) FAB score of 8.5 (3.5), considerably lower than that seen in the patients described in this thesis and their mean age was also greater (66.9 years). The mean FAB score in the 42 healthy controls studied was 17.3 (0.8) and in the PD subjects 15.9 (3.8). This compares favourably with the mean FAB score reported in an Italian study of 364 healthy controls subjects (Appollonio *et al.*, 2005).

The MSA-P patients in this thesis had a mean FAB score of 14.7 (2.5). PD patients had a FAB score of 16.5 (1.4). The value for the PD subjects is similar to that seen in the original report of the FAB (mean in PD=15.9, SD=3.8). Both of these values are close to the mean values seen in healthy control subjects but this should not be interpreted as showing intact cognitive functioning in PD as the sensitivity of the FAB to detect the mild executive deficits known to occur in PD (Emre, 2003) may be limited.

It is important to ascertain whether the FAB score, thought to be fairly specific for subcortical and executive dysfunction (Dubois *et al.,* 2000) is appropriately associated with more formal neuropsychological test scores and not excessively influenced by tests of general cognitive function. Dubois *et al* stated that the FAB score was not associated with the MMSE and hence concluded that the FAB score was not overly influenced by more general cognitive decline. In order to be certain of this, and to be certain that the FAB score is reflective of executive dysfunction, associations between the FAB and formal neuropsychological scores were studied.

5.5. Associations between the FAB and formal neuropsychology scores

5.5.1. Methods

Linear regression models were fitted to investigate the relationship between the FAB score and the test results from more formal tests of cognitive function.

Not all subjects were able to complete the full set of neuropsychological tests, (the maximum number of patients who had missing data on a particular cognitive test was five) and subjects with missing data were excluded from the statistical analysis for those tests.

To increase statistical power by ensuring that there were sufficient numbers of patients in each subgroup, we divided the whole sample of 40 patients into high or low FAB scores. The FAB cut-off score of 15 was used to separate the 40 cases into two groups: 21 with high ($\geq$15) versus 19 with low (<15) FAB scores. A series of t-tests were used to identify significant differences between the groups with high and low FAB scores. As we used a series of 38 t-tests, a Bonferroni correction was performed to adjust the level of significance for multiple comparisons. Only p values of p=0.001 or less were considered significant.

5.5.2. Results

Significant positive associations were found between the FAB total score and more formal tests of cognitive function across the entire patient cohort. These are shown in table 5.4. The associations between FAB scores and the Reitan Trail A, Recognition Memory for Faces, the Beck depression inventory, the Beck Anxiety Inventory, or the number of repetitions or intrusions on the verbal fluency test and the Rey AVLT were not significant.

Table 5.5 *Associations between the FAB and detailed neuropsychological assessment.*

Positive Correlations					
Tests influenced by frontal executive function			Tests of general cognitive function		
Test	*Correlation coefficient*	*P value*	*Test*	*Correlation coefficient*	*P value*
WAIS-R Similarities	0.68	< 0.001	NART	0.61	< 0.001
WCST correct categories	0.48	0.002	WAIS-R VIQ	0.71	< 0.001
Verbal Fluency phonemic	0.7	< 0.001	WAIS-R vocabulary	0.78	< 0.001
Verbal Fluency semantic	0.59	< 0.001	WAIS-R digit span	0.67	< 0.001
Verbal Fluency alternating semantic	0.75	< 0.001	MMSE	0.62	< 0.001
PASAT	0.65	< 0.001	Mattis DRS total	0.67	< 0.001
Mattis DRS Initiation/perseveration	0.6	< 0.001	Mattis DRS construction	0.44	< 0.001
Mattis DRS conceptualisation	0.39	0.01	Mattis DRS memory	0.36	0.03
			Mattis DRS attention	0.33	0.04
			Rey AVLT (various components)	0.39 - 0.65	0.01 - < 0.001
Negative Correlations					
Tests influenced by frontal executive function			Tests of general cognitive function		
Test	*Correlation coefficient*	*P value*	*Test*	*Correlation coefficient*	*P value*
WCST perseverative errors	-0.5	0.003	Cognitive decline NART- WAIS-R IQ	-0.41	0.017
WCST non-perseverative errors	-0.37	0.03			
Reitan TMT B	-0.65	< 0.001			
Reitan TMT B-A	-0.67	< 0.001			
Apathy Questionnaire	-0.42	0.009			

When studied separately with multiple regression analysis and co-varying for MMSE, Mattis DRS, age and disease duration, the FAB score in the PSP group alone, correlated with the following neuropsychological tests: PASAT ($R^2 = 0.78$, $t = 4.7$, $p = 0.001$), RTMT B ($R^2 = 0.93$, $t = -4.5$, $p = 0.02$), WCST perseverative errors ($R^2 = 0.7$, $t = -3.1$, $p = 0.02$) and the vocabulary sub-test of the WAIS-R ($R^2 = 0.7$, $t = 3.0$, $p = 0.01$). No correlations with other neuropsychological tests in the PSP group were detected.

There were no significant differences in UPDRS II, III or age between the patients with high versus low FAB scores ($p>0.05$). The patients with FAB scores ≥ 15 had a longer disease duration ($p=0.005$) but this did not reach significance after a Bonferroni correction (table 5.6). The group with low scores on the FAB (i.e. more severe executive dysfunction) had significantly lower scores on the Mattis DRS (total score and initiation subtest), the WAIS-R (Verbal IQ, vocabulary and similarities), Verbal fluency (phonemic, semantic and alternating semantic), the Paced Auditory Serial Addition Test and the 4^{th} and 5^{th} trials of the RAVLT.

No significant differences were detected between the groups on the Mattis DRS construction or memory subtests, measure of cognitive decline (difference between NART premorbid IQ estimate and WAIS-R current Verbal IQ), information retention across a 30 minute delay on the RAVLT (difference between delayed trial and trial 5), the number of errors or the percentage of perserverative errors on the WCST, the Reitan TMT-A, the recognition trial of the RAVLT, the recognition memory (RMT) for faces, scores on the Beck depression inventory or the Beck anxiety inventory.

The group with low scores on the FAB had a trend towards higher Reitan TMT-B, Reitan TMT B-A difference scores, number of perseverative errors on the WCST and scores on the Apathy Scale.

There was a trend towards lower scores on the NART VIQ, the Mattis DRS (attention and conceptualisation sub-tests), the MMSE, the second, third and delayed trial of the RAVLT, the number of categories correctly sorted, and the number of perseverative errors on the WCST. None of these reached significance after a Bonferroni correction.

Since the FAB includes one item on conceptualization/similarities and one verbal fluency item, we also examined the correlation between adjusted FAB scores

(omitting these items) and the Similarities subtest of the WAIS-R, the conceptualisation subtest of the Mattis DRS and the phonemic verbal fluency scores. The Similarities subtest of the WAIS-R ($r=0.49$, $p=0.003$) and the phonemic verbal fluency test ($r=0.69$, $p<0.001$) had positive and significant correlations with the adjusted FAB score, whereas the conceptualisation subtest of the Mattis DRS did not.

With the exception of the phonemic verbal fluency, WAIS-R Similarities and DRS conceptualisation subtests, which have overlapping items in the FAB, the variables that showed sizeable and significant correlations with the FAB score were entered in a stepwise regression analysis to determine which of them accounted for the greatest proportion of its variance. These were duration of illness, the NART estimate of premorbid IQ, WAIS-R, Verbal IQ, MMSE, Mattis DRS total score, RAVLT trials 1-5 and delayed recall, WCST categories correct, perseverative and non-perseverative errors, semantic verbal fluency, alternating semantic fluency, PASAT, Reitan TMT B, the TMT B-A difference score, the Apathy scale, the measure of cognitive decline (difference between NART and WAIS-R Verbal IQ). Only the alternating semantic fluency entered the stepwise regression and accounted for 80% of the variance of the FAB scores. (Standardized coefficient = 0.906, Adjusted R square 0.904, $F(1,11)= 50.2$, $p<0.001$).

Table 5.6 *Mean (SD) scores for age, disease duration, UPDRS and neuropsychological tests for patients with FAB scores equal to or higher than 15 or those with FAB scores lower than 15. Significant differences after a Bonferroni correction are in bold.*

	FAB score ≥ 15			FAB score < 15			T-test
	n	Mean	SD	n	mean	SD	P value
Patient details							
Age (years)	21	62.7	6.6	19	65.8	8	=0.2
Disease Duration (years)	21	9.5	6.7	19	4.8	2	=0.005
UPDRS II	21	18.9	7.5	19	20.3	7.5	=0.5
UPDRS III	21	21.2	8	19	21.1	8.9	=0.9
Tests of general cognitive function or mood							
MMSE	21	27.9	2.2	19	25.1	3	=0.002
NART VIQ	20	117.8	12.1	15	107.9	12.7	=0.03
WAIS-R VIQ	20	113.2	16.7	18	93.6	15.3	**=0.001**
Vocabulary (scaled)	19	13.5	3.6	17	8.8	3.0	**<0.001**
Cognitive decline (NART- WAIS-R IQ)	19	3.4	12.8	14	11.9	11.7	=0.06
Mattis DRS total	21	137.6	6.7	18	121.8	18.2	**=0.001**
Mattis DRS memory	21	23.0	3.9	18	21.8	3.7	=0.3
Mattis DRS Construction	21	5.5	1.3	18	4.8	2.1	=0.2
REY1	21	5.5	2.7	19	4.4	2.2	=0.2
REY2	21	8	2.8	19	5.9	2.8	=0.02
REY3	21	9.9	2.7	19	7.3	3.5	=0.01
REY4	21	11.2	2.2	19	8.0	3.3	**=0.001**
REY5	21	11.7	2.7	19	8.5	3.1	**=0.001**
REY Delay	21	10.4	4.7	19	6.3	4.4	=0.007
Memory Loss with delay (REY 5 – REY Delayed recall)	21	1.3	3.2	19	2.2	2.2	=0.3
Recognition Memory for Faces	21	22.1	2.8	15	21.1	3.9	=0.4
Beck Depression Inventory	21	11.1	8	18	13.6	6.3	=0.3
Beck Anxiety Inventory	21	16.0	12.9	18	11.7	8.2	=0.3
Tests of executive function							
Mattis DRS Initiation	21	34.3	3.2	19	25.9	7.9	**<0.001**
Mattis DRS Concenptualization	21	37.6	1.7	19	32.2	9.4	=0.02
WAIS Similarities (scaled)	18	12.1	3.7	17	7.8	3.3	**=0.001**
Mattis DRS attention	21	37.8	0.5	19	33.5	6	=0.03
Wisconsin CST categories sorted	21	4.3	2.5	18	2.7	2.2	=0.04
Wisconsin CST number of errors	19	6.7	7	14	11.6	8.4	=0.07
Wisconsin CST perseverative errors	19	3.6	5.1	14	6.8	5	=0.03
Wisconsin CST % perseverative errors	19	28.3	27.4	14	38.1	18.7	=0.3
Reitan A	20	76.5	50.9	13	104.6	82.1	=0.2
Reitan B	19	136.4	65.1	10	227.3	106.3	=0.03
Reitan Trail B-A difference	19	69.4	48.5	10	156.8	88.1	=0.01
Verbal fluency phonemic (total FAS)	21	41.5	16.2	18	16.4	9.1	**<0.001**
Verbal fluency semantic animals	21	14.3	5	18	8.7	4.5	**=0.001**
Verbal fluency alternating categories	21	15.8	4.4	18	7.7	4.7	**<0.001**
PASAT	20	20.2	6.7	16	13.6	5.8	**=0.001**
APATHY scale score	20	11.8	6.5	18	21.8	12	=0.004
REY Repetition errors	21	13.7	1.6	19	12.5	2.9	=0.1
REY Intrusion errors	21	0.6	0.9	19	1.0	1.4	=0.3

5.5.3. Conclusions

The FAB showed sizeable and significant correlations in the expected direction with tests of executive function (verbal fluency, Reitan TMT, Similarities, WCST, PASAT, the initiation and conceptualisation subtests of the Mattis DRS, and the apathy questionnaire). FAB scores also showed significant correlations with more general tests of cognitive function such as the MMSE, Mattis DRS total score, Verbal IQ, NART estimated IQ and measures of cognitive decline.

Patients with high FAB scores significantly differed in terms of scores on tests of executive 'frontal' function namely the Mattis DRS initiation and perseveration subtest, verbal fluency, the PASAT and WAIS Similarities as well as a trend towards higher scores on the apathy questionnaire than those with low FAB scores.

The associations between the FAB and the age of the patient and severity of illness assessed on the UPDRS II and III subsets were neither large nor significant. A significant association with duration of illness ($r=0.41$, $p=0.009$) arose because the PD subjects had significantly higher FAB scores and longer disease duration. This occurred simply because all subjects considered together and should not be interpreted as suggesting that FAB scores increase as bradykinesia and rigidity increase.

The association between the FAB score and the NART estimated pre-morbid IQ suggests that pre-morbid intellectual functioning might influence the FAB score. This has also been suggested in a large study of the FAB score in healthy control subjects (Appollonio *et al.*, 2005). Hence education or pre-morbid IQ should be taken into account when interpreting the results of the FAB score.

Together with disinhibition and impulsivity, apathy is considered to be one of the key behavioural manifestations arising from frontal dysfunction (Marin, 1996). It is

interesting to note that while FAB scores were significantly correlated with scores on the Apathy scale, the correlations between FAB scores and measures of depression or anxiety were not significant.

One advantage of the FAB for assessing cognitive function in PSP is that it does not contain any subtests dependent on intact oculomotor function. The gaze palsy in PSP should not therefore contribute to the difference between FAB scores in the PSP, MSA-P and PD patients. This is not the case for the more formal cognitive tests and particularly scores on the WCST, the Reitan TMT and the recognition memory test for faces which can all be influenced by the oculomotor dysfunction in PSP. However this should not affect the results in this study as all patients who were unable to complete these tests due to severe gaze palsy were excluded from the analysis for these specific tests.

These results are partly in agreement with the findings of Dubois et al in that the FAB score reflects scores on more formal tests of executive function (Dubois *et al.*, 2000). However, in addition, the results outlined in this chapter demonstrate that the FAB may help differentiate MSA-P and PD and Dubois *et al* may have failed to show this because of their small MSA sample size (n=6).

These results extend those of Dubois *et al* by showing clear correlations between the FAB and not only the WCST but also other formal tests of executive function. In contrast to the results of Dubois *et al*, in this study, FAB scores also showed significant correlations with more global measures of cognitive function such as the MMSE, as well as the Mattis DRS total score (although this test has components that test executive function), the WAIS-R Verbal IQ, and the estimate of premorbid IQ from the NART. Furthermore, there were significant differences in the mean scores of

several tests of general cognitive function between the subgroups with high and low FAB scores.

This is not surprising given that degenerative disorders often have multiple cognitive domains affected simultaneously. This means there is an inevitable element of co-linearity with frontal and global cognitive decline and suggests that global cognitive decline, as well as sub-cortical dementia is likely to affect the FAB score.

The correlation between the NART and the FAB score could suggest that a lower pre-morbid IQ is associated with poorer performance on the FAB. However there was no significant difference between the NART scores in the high and the low FAB groups.

A criticism of other bedside tests of frontal function such as the EXIT 25 test, has been that they showed significant correlations with more global measures of cognitive function such as the MMSE, which meant they measured aspects of cognition other than executive function (Royall *et al.*, 1992). The results above suggest that this is also true of the FAB. Nevertheless, given its discriminant validity and its ability to correctly differentiate patients with PSP, MSA-P or PD, administering the FAB at the bedside may be useful in the routine clinical assessment of bradykinetic rigid syndromes.

That simple and alternating semantic and phonemic verbal fluency tests have been shown to correctly differentiate PSP, PD and MSA and are considered sensitive measures in the differential diagnosis of these disorders (Lange *et al.*, 2003) is reflected in the results of this thesis in which lexical fluency seems to be one of the more important sub-scores of the FAB with regard to its ability to differentiate PSP, PD and MSA-P.

Studies measuring reaction times and event-related potentials have shown that speed of processing is significantly slower in PSP than in PD and in PSP than in MSA or PD (Dubois *et al.*, 1988; Pirtosek *et al.*, 2001). Hence a potential confounding factor when applying the FAB to these parkinsonian syndromes is that processing speed might have been slower in PSP than in MSA-P or PD and this could have influenced the group differences in FAB scores. However, with the exception of the lexical fluency subtest, none of the FAB subtests have a time limit. For this reason, one might suggest that the significant differences between the PSP, MSA and PD patients are not simply reflective of differences in processing speed. However lexical fluency is one of the most important FAB subscores with regard to correct classification of PSP, MSA and PD, and a measure of motor speed (Reitan TMT A) did not correlate with FAB scores, while the more cognitively demanding Reitan TMT B, and the TMT B-A difference score did. These psychological tests are not restricted to testing motor, cognitive or processing speed independently of each other, but could be interpreted as suggesting that speed of cognitive processing may be relevant to performance on the FAB in bradykinetic rigid syndromes after all.

5.6. Overall conclusions

These results confirm that PSP has more severe subcortical and executive dysfunction than MSA-P and PD. The lack of any significant difference between the patient groups in the rate of cognitive decline is likely to be a consequence of the short interval between testing, as well as patient drop out.

FAB scores at the initial assessment discriminated PSP from MSA-P and PD and also MSA-P from PD and appear to have clinical validity as a test supportive of the diagnosis of PSP. While a normal or abnormal FAB score cannot be defined because the FAB has not been administered to a healthy control group here, a score of <15

indicates significant impairment, well below the mean (-1 SD) seen in control groups in two studies (Appollonio *et al.,* 2005; Dubois *et al.,* 2000).

The FAB scores in all patients were associated predominantly with performance on tests of executive function, further supporting the use of the FAB as a quick bedside test in differentiating the more severe executive dysfunction of PSP from that seen in MSA and PD. However the fact that in contrast to previous studies, significant correlations with tests of general cognitive were detected, suggests that the FAB score is influenced by more global cognitive impairment and the results of the test should be interpreted cautiously and in the correct clinical context.

Imaging correlates of these findings in PSP, have been examined and are discussed in later chapters.

6. Cross Sectional Volumetric Imaging

6.1. Introduction

The phenotypic differences that characterise different neurodegenerative diseases are underpinned by the topography of cerebral dysfunction. Certain diseases have a predilection for affecting particular regions of the brain, and the resulting atrophy may be visible during life when the brain is imaged. In PSP, a reduction in midbrain size is the imaging finding most frequently quoted. Qualitative assessments regarding its size and shape in PSP mention concavity of the superior midbrain in the mid-sagittal view on MRI (Righini *et al.*, 2004), and have drawn comparisons to the profile of a hummingbird (Kato *et al.*, 2003) and on axial views to the "morning glory" flower (Adachi *et al.*, 2004). Volume differences in numerous other regions have been commented upon including the third ventricle (Savoiardo *et al.*, 1994; Schrag *et al.*, 2000), the corpus callosum (Yamauchi *et al.*, 2000) and the frontal and temporal lobes (Schrag *et al.*, 2000). Many of these studies have relied on qualitative judgements as to the presence or absence of atrophy, although more recently, studies based on quantitative volumetric and area measurements have been published. Further studies, however, are needed to validate published findings. It is also important to study regions quantitatively, such as the third ventricle, in which only qualitative data is so far available (Schrag *et al.*, 2000). Quantitatively studying brainstem regions such as the midbrain, pons and cerebellum and analysing clinico-radiological correlations in PSP and MSA-P may further understanding of the clinical phenotype in these diseases.

Region of interest (ROI) segmentation requires assumptions regarding the areas we expect to be affected in PSP. These assumptions are largely based on: (1) information

from previous imaging studies in PSP; (2) Information from post-mortem histopathological studies; (3) assumptions regarding which anatomical regions are expected to be involved given the clinical phenotype of a disease. In this way, ROI studies are at least partly hypothesis driven.

Manual region of interest studies remain the gold standard for detection of differences in brain and regional brain volume between diseases and healthy controls. However, novel image analysis techniques such as voxel based morphometry (VBM, Chapter 7) can be used to identify regional differences and support the findings of hypothesis driven ROI studies, without making any a priori assumptions regarding which regions to study.

Segmentation is the process of outlining a particular region of interest (ROI) on an MRI scan. Outlining of large CSF spaces is straightforward as their borders are often easily identified. Likewise, outlining the borders of a structure surrounded by a CSF space, such as the anterior border of the pons, is aided by that high contrast boundary. However CSF may not surround all parts of a ROI, and this makes segmentation more difficult and requires subjective and often arbitrary assessment of the boundaries concerned. For small structures, that are difficult to outline, variability secondary to measurement error can mask true group differences in regional volume. Because of this, clear protocols for segmentation need to be adhered to in order to minimise error and to standardise the anatomical cut-offs in regional volume measurement. This can mean that some of the relevant structure volume is excluded. For regions that have not been studied previously, preliminary region of interest studies can be analysed and segmentation protocols adapted and refined in order to minimise measurement error.

For all image segmentation performed in this study, the digitised structural MR images were transferred to a SUN workstation (Sun Microsystems Inc, Mountainview, CA) and image analysis undertaken using the MIDAS software package (Freeborough *et al.,* 1997) with images registered to standard space. The MIDAS package, allows manual segmentation of regions of interest using a mouse-driven cursor. The simultaneous display of orthogonal views allows the operator to outline the structure in the coronal, axial or sagittal view whilst the segmentation is updated in real time in another plane, viewed in a second window. This aids in decisions about where regional boundaries are defined. Segmentation protocols are described in more detail in Chapter 6.5.

6.2. Aims

The aims of cross sectional volumetric image analysis were directed principally towards: 1) identifying any novel regions of volume difference that may help to discriminate PSP during life; 2) identifying group differences in whole brain volumes between PSP, MSA-P, PD and healthy controls; and 3) undertaking region of interest segmentation in order to quantify regional, particularly brainstem atrophy patterns using manual, semi-automated and template-based ROI analysis techniques in these diseases; and 4) determining the combination of regional atrophy that best differentiates PSP and MSA-P.

In order to achieve these aims, new region of interest measurement protocols had to be devised and assessed for repeatability and reliability.

6.3. The superior cerebellar peduncle as a potential discriminator of PSP

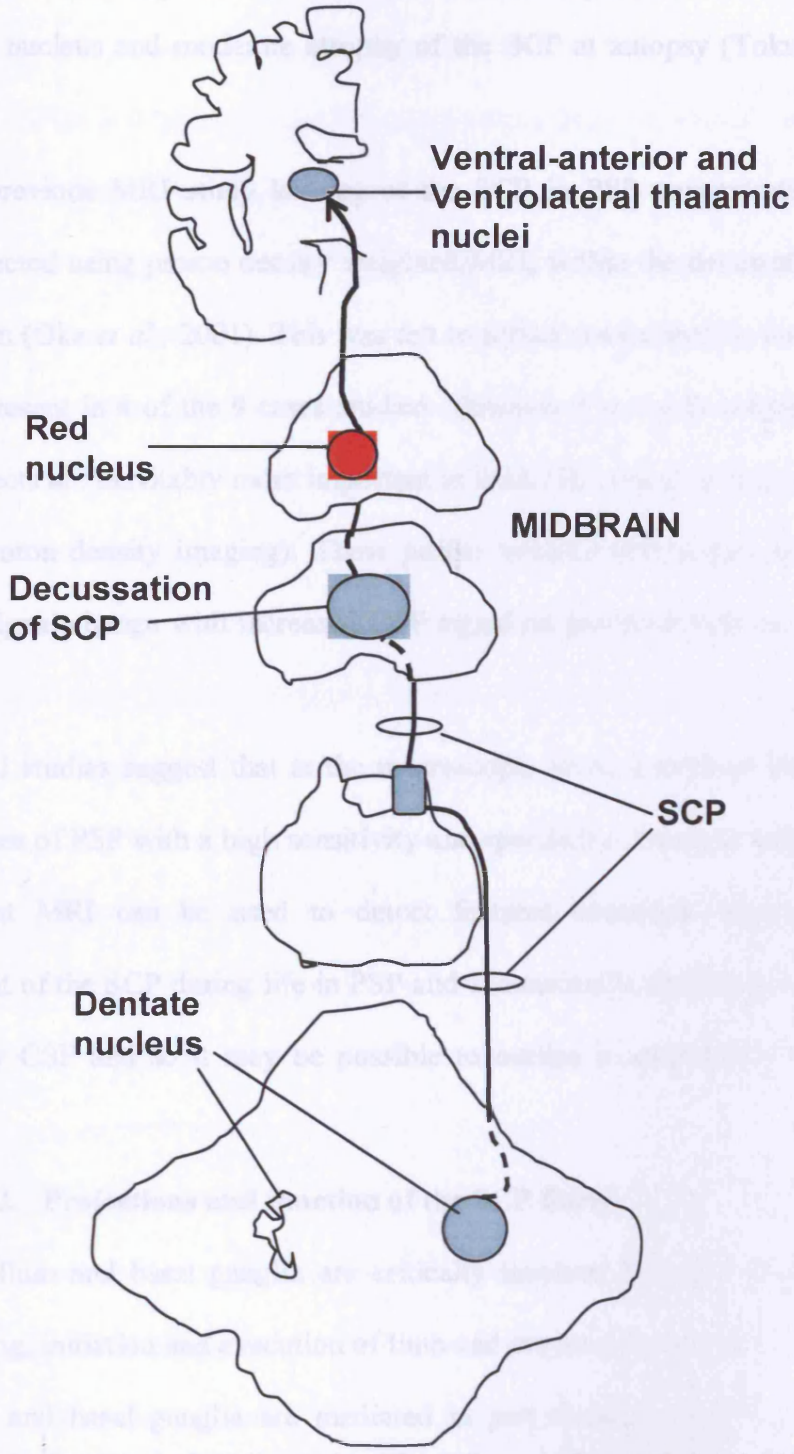
6.3.1. Introduction

In their original description of PSP in 1964 (Steele *et al.,* 1964), the authors described cell loss and gliosis in the cerebellar dentate nucleus as well as atrophy and

demyelination in the superior cerebellar peduncle (SCP) or brachium conjunctivum. This structure contains fibres projecting from the dentate nucleus to the ventral lateral nucleus of the thalamus. Recently, macroscopic atrophy of the SCP has been systematically studied pathologically in PSP and neurodegenerative disease control subjects, concluding that SCP atrophy at post mortem reliably differentiates between PSP and other clinically related neurodegenerative diseases (Tsuboi *et al.*, 2003).

The SCP consists mainly of efferent fibres from the cerebellar dentate nucleus, which run in the lateral wall of the fourth ventricle. The fibres enter the posterior aspect of the pontine tegmentum, postero-medial to the upper fibres of the middle cerebellar peduncle, inferior to the inferior colliculus. Just anterior to the SCP fibres in this part of the brainstem are the fibres of the mesencephalic tract of the trigeminal nerve and the motor nucleus of the Trigeminal nerve. The SCP fibres continue in the upper pons, move medially and begin to decussate as the pons merges with the midbrain. Within the midbrain the fibres are situated between the interpeduncular fossa anteriorly and the IVth ventricle posteriorly and are bounded by the pontine reticular formation laterally. Within the midbrain, the SCP fibres sit anterior to the cerebral aqueduct and the peri-aqueductal grey (PAG), the medial longitudinal fasciculus and tectospinal tract. The interpeduncular nucleus and fossa are anterior to the SCP with the substantia nigra and corticopontine fibres anterolateral. The fibres enter the cerebellorubrothalamic tract, which is sited between the red nucleus posteromedially and the substantia nigra anterolaterally. The *in-vivo* anatomical mapping of these brainstem axonal connections has been demonstrated recently, using diffusion tensor MRI (Mamata *et al.*, 2002; Stieltjes *et al.*, 2001). A schematic representation of the pathway is shown in figure 6.1.

Figure 6.1 *Schematic diagram of the cerebello-rubrothalamic pathway*



The schematic diagram displays the origin of the pathway in the cerebellar dentate nucleus from which the fibres run in the superior cerebellar peduncle, decussating in the midbrain above the level of the superior colliculi. Some of the fibres synapse in the red nucleus, others continue rostrally to synapse in the ventral anterior and ventrolateral thalamic nuclei.

Atrophy of the SCP on MRI has been noted, in SCA-3 and SCA-6 (Murata *et al.*, 1998b; Murata *et al.*, 1998a), and in one case of SCA-3 coming to post mortem,

atrophy of the SCP seen on MRI in life was confirmed by grumose degeneration of the dentate nucleus and moderate atrophy of the SCP at autopsy (Tokumaru *et al.*, 2003).

The only previous MRI study looking at the SCP in PSP concentrated on signal change detected using proton density weighted MRI, within the decussating fibres in the midbrain (Oka *et al.*, 2001). This was felt to reflect demyelination and gliosis but was only present in 4 of the 9 cases studied. However it is worth noting that partial volume effects are inevitably more important in thick slice imaging (e.g. 3-5mm axial slices in proton density imaging). These partial volume effects may make atrophy appear as signal change with increased CSF signal on proton density or T2 "shining through".

Pathological studies suggest that at the macroscopic level, a reduced SCP size may identify cases of PSP with a high sensitivity and specificity. Previous imaging studies suggest that MRI can be used to detect features consistent with pathological involvement of the SCP during life in PSP and anatomically, the structure is partially bounded by CSF and so it may be possible to outline it accurately on volumetric scans.

6.3.2. Projections and function of the SCP fibres

The cerebellum and basal ganglia are critically involved in motor control and the programming, initiation and execution of limb and eye movements. The outputs of the cerebellum and basal ganglia are mediated in part through projections to various thalamic nuclei. Dysfunction of these pathways results in a number of profound motor disturbances.

The projection of the dentate nucleus to different cortical areas has been disputed (Asanuma *et al.*, 1983; Holsapple *et al.*, 1991). More recently, it has generally been

accepted that the dentate nucleus contains a topographical motor map of the body, with face, arm and leg areas projecting via the thalamus to the motor cortex (Hoover and Strick, 1999). Cells projecting to the motor cortex make up only about 30% of the dentate nuclear volume, so 70% of its cells must project elsewhere.

The nucleus probably contains multiple topographically organised body maps projecting to different cortical areas and more recently, it has been shown that cells project to the pre-frontal and posterior parietal cortex as well as to the primary motor cortex (Dum and Strick, 2003). The non-motor domain in the ventral portion of the dentate contains output channels concerned with aspects of cognition and visuo-spatial function. These ventral portions are activated during a variety of tasks involving short term working memory, rule based learning and higher executive functions such as planning (Kim *et al.*, 1994).

The functions of the neurons travelling within the SCP are complex and varied. While the majority of the fibres originate from the dentate nucleus, a small number come from the nucleus interpositus and the fastigial nucleus. Lesional studies often provide good information regarding the function of specific structures and this has certainly been true of a large retrospective study of basal ganglia lesions (Bhatia and Marsden, 1994). Lesions isolated to the superior cerebellar peduncle have been reported as causing impaired smooth pursuit while saccadic eye movements were preserved (Ohtsuka *et al.*, 1992). Other authors have reported that a lesion of the dentate and its projections results in severe postural tremor. This is not a typical clinical finding in PSP, but can certainly occur. A lesion in the superior cerebellar peduncle, at the level where it is closest to the oculomotor nerve fascicles at the caudal end of the red nucleus has been cited as the most likely site causing Claude's syndrome (Seo *et al.*,

2001). This syndrome is characterized by ipsilateral occulomotor nerve palsy and contralateral cerebellar ataxia. Again these are not typical clinical features of PSP.

Lesional studies in animals have shown that damage to the SCP results in abolition of eyelid and limb conditioned responses (McCormick *et al.*, 1982; Voneida, 2000). It is not immediately apparent that we can use this information to postulate how pathology in the dentate nucleus and subsequent atrophy in the SCP may be correlated with loss of specific functions and clinical signs in PSP, particularly as the pathology in PSP is not confined to this one region.

6.3.3. Aims

The purpose of regional segmentation of the SCP on volumetric T1 weighted MRI, was to determine whether the MRI appearances of this structure can help to distinguish PSP from other Parkinsonian syndromes in life, and whether this can be specifically related to particular clinical features of the disease. Prior to undertaking this analysis however, a method of segmenting the SCP and calculating its volume needed to be devised.

The aims of this preliminary study of the SCP were:

1. To devise a protocol for segmentation of the SCP.
2. To assess the repeatability and reliability of the protocol in a preliminary series of segmentations.
3. To assess the methodology looking for error between paired segmentations and to alter the segmentation protocol as required.
4. To repeat the segmentations and reassess repeatability and reliability.
5. With the final protocol, assess the ability of SCP volume to differentiate PSP from other degenerative diseases.
6. To review the clinical value of SCP atrophy as a marker for PSP.

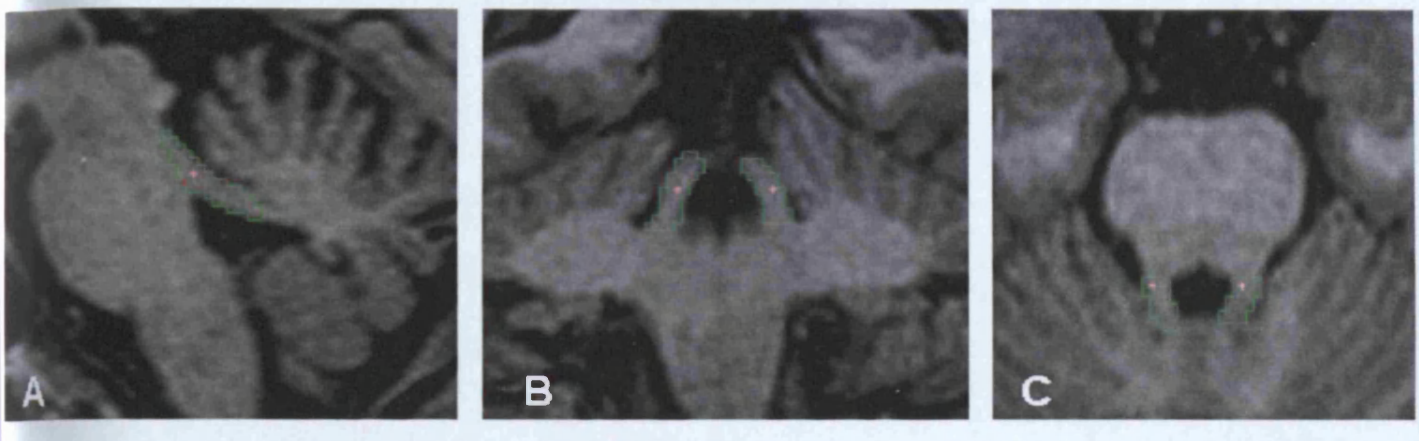
6.3.4. Preliminary methods of SCP segmentation

A preliminary study of the SCP was undertaken once a suitable number of initial MRI scans had been acquired. T1 volumetric MRI scans in 12 patients with PSP, 12 with PD, 8 patients with MSA-P and 11 non-diseased controls were included. MRI scans were acquired on the same 1.5GE Signa Unit (General Electric, Milwaukee, WI) using the T1 acquisition parameters described previously. The images were transferred to a SUN Microsystems workstation for analysis.

The images were acquired coronally and the SCP was initially segmented in this plane, from anterior to posterior. The superior margin was taken at the level of the inferior colliculus and the inferior margin the parabrachial sulcus. The SCP is bounded medially and laterally by CSF in coronal section. Sagittal and axial views were available for comparison throughout and subsequently the axial sections were checked and the segmentations modified. On Sagittal section, the anterior border was taken as where the SCP entered the brainstem and the posterior border, where it joins the cerebellum (figure 6.2A). This measurement protocol only takes into account the fibres of the SCP after they have left the cerebellum and prior to entry into the brainstem.

The images were segmented in standard space, in random order by a single rater (myself) blinded to the identity and clinical diagnosis in each case. A second blinded measurement was then made once the complete data set had been analysed and compared to the first set of measurements to allow assessment of the repeatability and reliability of the method.

Figure 6.2 *Examples of superior cerebellar peduncle segmentation.*



T1 weighted volumetric MRI of the superior cerebellar peduncle in a healthy control subject demonstrating outlining of the structure (**A**- sagittal, **B**- coronal, **C**- axial views of MIDAS window, showing the superior cerebellar peduncle-SCP, outlined in green). The superior cerebellar peduncle contains efferent fibres from the deep nuclei of the cerebellum (mainly the dentate nucleus) which pass superiorly to the red nucleus and thalamus, decussating in the caudal midbrain.

6.3.5. Repeatability and reliability measures

When an instrument is used to measure a clinical quantity it is desirable that pairs of measurements made on the same individual using the same instrument give "similar" results in the absence of any other variation in circumstances. Repeatability and reliability are both concerned with quantifying the degree of similarity between such pairs of measurements.

Repeatability relates to the standard deviation of within person differences and hence is measured in the same units as the clinical quantity of interest. Reliability relates the variation seen within individuals, to that seen between individuals and hence is unit free.

If a second instrument is used to measure the same clinical quantity, then interest centres on the agreement between the two instruments.

In deciding whether two instruments agree, consideration needs to be given to bias (does one instrument give higher measurements than another) and variability (how

variable are the differences between measurements made on the same individual using different instruments).

If a variable is measured twice in n individuals using a single measurement technique, then the within subject standard deviation (σ_w) can be used as a measure of repeatability. The smaller the within subject standard deviation, the greater the repeatability. Alternate measures of repeatability, such as the standard deviation of differences and the 95% reference range for differences (i.e for a variable, 95% of observations are expected to lie within ± 1.96 standard deviations from the mean) can also be used.

For a measurement technique with the best possible repeatability, observed differences should have zero mean and constant variance. The mean difference in measurements should be seen to lie at around zero.

In order to estimate reliability, the simple model for repeatability can be extended to give the between subject variability. This can be estimated using analysis of variance. The estimated between person variance is given by:

$$\sigma_b^2 = (\text{BMS-RMS})/m$$

where m is the number of measurements per subject, BMS is the between subject mean squared difference and RMS the residual or within individuals mean squared difference.

The reliability coefficient can be calculated from:

$$\frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}$$

The reliability coefficient is the proportion of the total variance that is due to between person variability and as such takes values in the range (0-1) and is unit free. The

closer it is to 1, the more reliable is the method. Another name for the reliability coefficient is the intraclass correlation coefficient (ICC).

6.3.6. Results

The results of the mean difference in blinded segmentation of the SCP and the calculated within person SD and ICC in the preliminary measurements are detailed below.

Table 6.1 *Superior cerebellar peduncle repeatability and reliability measures*

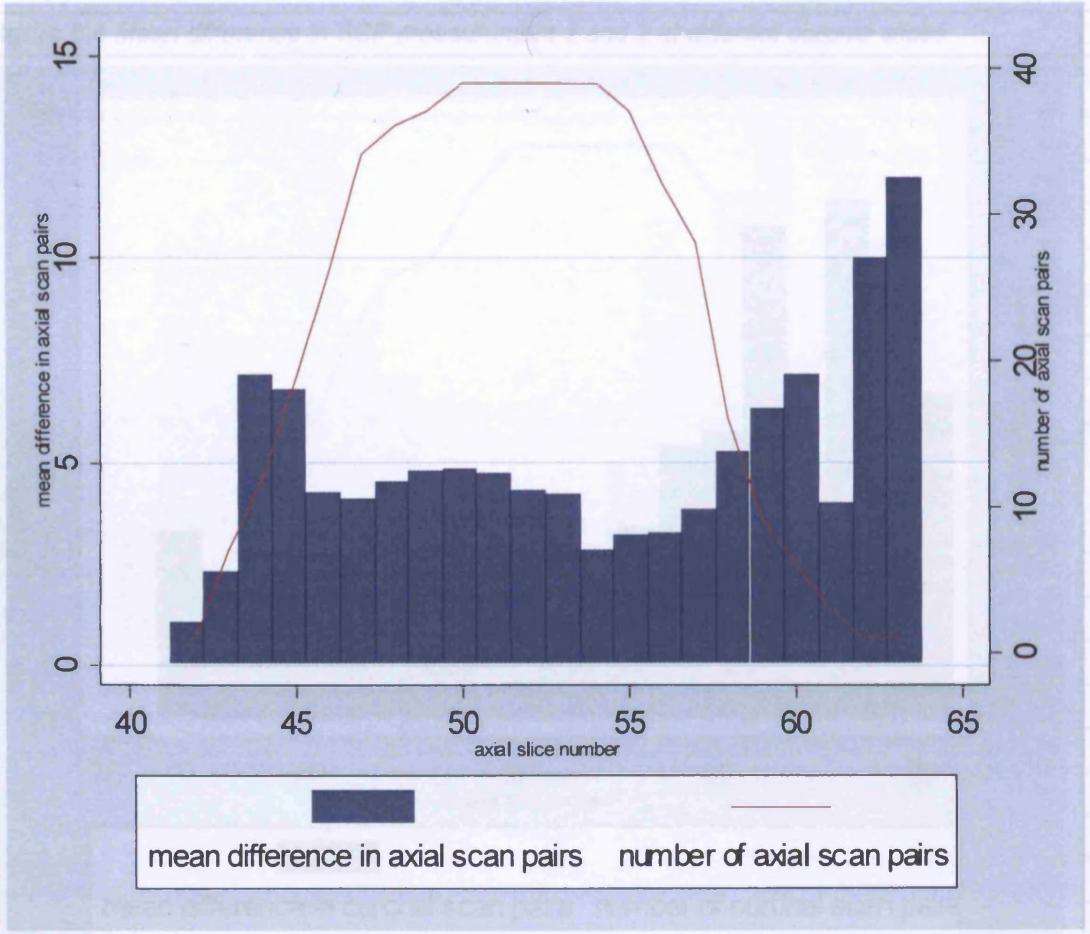
	Pairs of observations	Mean (SD) SCP volume (ml)	Range (ml)	95% C.I	Mean absolute difference in segmentations	Within person sd (ml)
Initial protocol	43	0.312 (0.039)	0.176 – 0.462	0.295-0.329	0.02	0.031
Intraclass correlation coefficient – 0.850 (95% C.I- 0.77-0.93)						
SCP- superior cerebellar peduncle; sd-standard deviation; C.I-confidence interval						

From these data, the method seems repeatable (within person standard deviation- 0.031ml) and reliable (ICC – 0.850). However whether this can be improved upon by revising the methodology depends upon determining where the differences between an individuals first and second assessment occurred, and if altering the technique can minimise these errors without creating new ones.

6.3.6.1. Where do measurement inconsistencies arise?

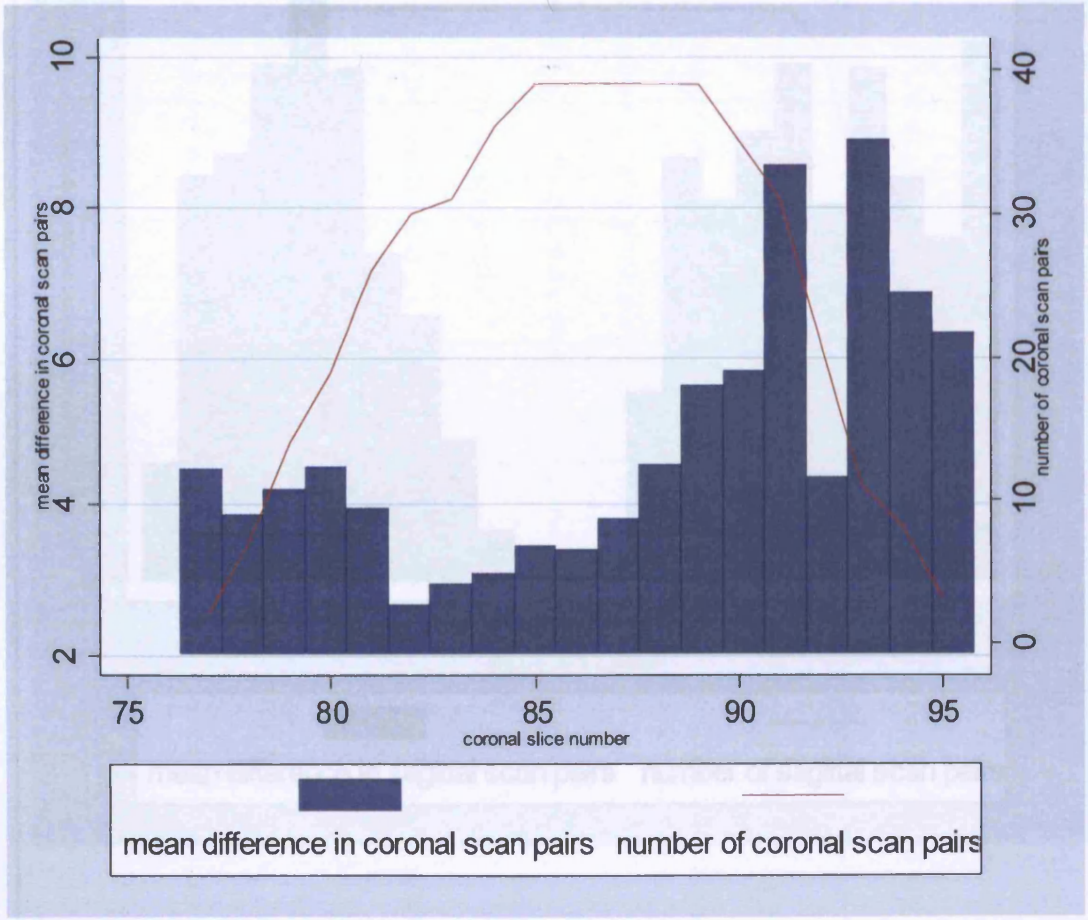
For segmentations in the axial, coronal and sagittal planes, the area segmented on each scan slice, in standard space, was recorded. The difference in the first and the second blinded measurements at each level was noted together with the total number of paired scans recorded at that scan slice. From this it is possible to assess where the greatest disparity between the first and second measurements occurred. The first and the second blinded segmentations were reviewed manually for each individual regardless of diagnosis and the location of the most obvious disparities between the first and the second segmentations was noted.

Figure 6.3 *Mean difference in SCP measurement 1 and 2 at different axial slices*



Axial scan data		
Slice number	**Mean difference in area**	**n**
42	1.00	1
43	2.29	7
44	7.17	12
45	6.79	19
46	4.23	26
47	4.12	34
48	4.50	36
49	4.81	37
50	4.87	39
51	4.72	39
52	4.31	39
53	4.21	39
54	2.82	39
55	3.22	37
56	3.25	32
57	3.82	28
58	5.25	16
59	6.33	9
60	7.17	6
61	4.00	3
62	10.00	1
63	12.00	1
n-number of measurements recorded at that axial scan level		

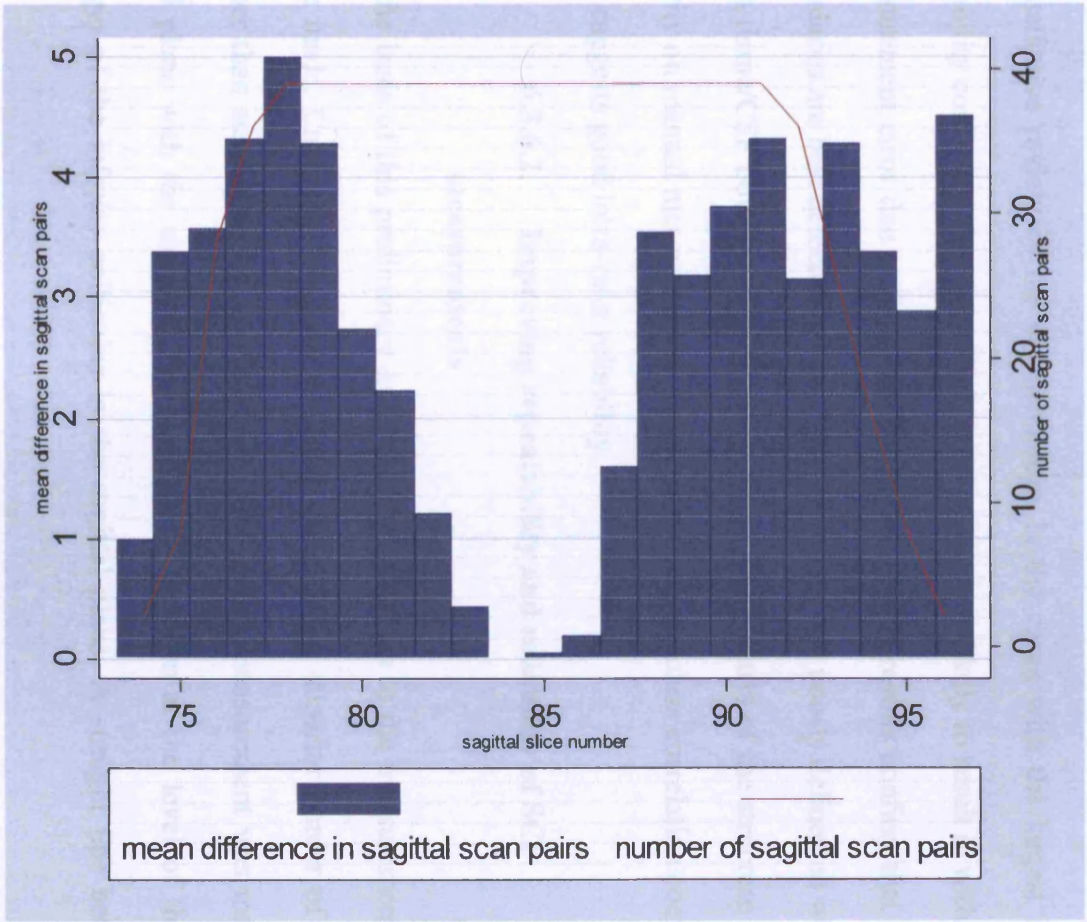
Figure 6.4 *Mean difference in SCP measurement 1 and 2 at different coronal slices*



Coronal scan data		
Slice number	**Mean difference in area**	**n**
77	4.50	2
78	3.86	7
79	4.21	14
80	4.53	19
81	3.96	26
82	2.67	30
83	2.94	31
84	3.08	36
85	3.44	39
86	3.38	39
87	3.82	39
88	4.56	39
89	5.62	39
90	5.83	35
91	8.55	31
92	4.38	21
93	8.91	11
94	6.88	8
95	6.33	3

n-number of measurements recorded at that coronal scan level

Figure 6.5 *Mean difference in SCP measurement 1 and 2 at different sagittal slices*



Sagittal scan data		
Slice number	**Mean difference in area**	**n**
74	1.00	2
75	3.38	8
76	3.57	28
77	4.31	36
78	5.00	39
79	4.28	39
80	2.74	39
81	2.23	39
82	1.21	39
83	0.44	39
84	0.00	39
85	0.05	39
86	0.18	39
87	1.59	39
88	3.54	39
89	3.18	39
90	3.74	39
91	4.31	39
92	3.14	36
93	4.27	26
94	3.38	16
95	2.88	8
96	4.50	2

n-number of measurements recorded at that sagittal scan level

Reviewing the results of analysis of the initial and repeated blinded measurements (figures 6.3-6.5), it is apparent from the tabulated and the graphical data, where the most important discrepancies between paired measurements lie. On the axial measurements, the most significant errors were made superiorly, between slices 42 and 46 and inferiorly between slices 60-65 (figure 6.3). Coronally, the biggest discrepancies arose anteriorly, between slices 90 to 95 with a few less significant errors posteriorly (figure 6.4). On the sagittal measurements, the greatest errors were made laterally (figure 6.5). However in all planes, in scan slices where the mean difference in measurement one and two was large, the actual number of subjects with scans that included tissue in that slice was small and unlikely to be of much significance at a group level.

A qualitative review of the segmentations in the cases with the largest intra-rater variability confirmed that these regions were the most likely to result in within subject measurement error due to the protocol utilised. These results confirm that the tissue boundaries are best detected at brain/CSF interfaces and poorly delineated where clear brain tissue/CSF boundaries were not present. These parts of the structure rely most heavily on manual measurements. Despite this, an interclass correlation coefficient of 0.85 suggests good intra-rater reliability.

6.3.6.2. Improving repeatability and reliability of SCP measurements

On the basis of this preliminary analysis, modifications to the measurement protocol were made. Care was taken in the axial plane at the superior border of the SCP. Rather than starting in the coronal plane, the initial measurement was made in the axial plane with the superior border of the SCP taken at the level of the inferior margin of the inferior colliculus in the sagittal plane. A straight line between the

inferior border of the inferior colliculus and the posterior border of the pons marked the anterior border of the SCP. The posterior border of the SCP was marked at a point where it entered the cerebellum and became indistinguishable from the cerebellar hemisphere. The axial sections were carefully checked with the sagittal and coronal sections and any deviation from the protocol rectified appropriately. Particular care was taken in the slices prone to measurement inconsistencies. As before, the inferior border of the SCP was taken as the level of the parabrachial sulcus. Repeatability and reliability measures were undertaken as before and are shown in table 6.2.

Table 6.2 *Superior cerebellar peduncle repeatability and reliability measures for modified protocol*

	Pairs of observations	Mean (SD) SCP volume (ml)	Range (ml)	95% C.I	Mean difference in segmentations	Within person sd (ml)
Initial protocol	43	0.273 (0.063)	0.164-0.395	0.257-0.289	-0.054	0.058
Interclass correlation coefficient – 0.215 (95% C.I- 0 - 0.50)						
SCP- superior cerebellar peduncle; sd-standard deviation; C.I-confidence interval						

Clearly using the revised method did not result in improved reliability (ICC) or repeatability of the SCP segmentation.

6.3.7. Conclusions

In conclusion, the ICC and within person standard deviation confirm that the first method for segmentation of the SCP is both reliable and repeatable. Using the revised method did not improve on the original method. The mean SCP volume using the revised method is slightly lower which would be expected given that the revised technique used a more rigid cut off for the anterior border of the SCP. Therefore for the purposes of SCP segmentation on the entire data set, the original method was used. The results of SCP volume analysis on the entire data set are reported in Chapter 6.5.4. The clinical utility of these measurements is commented upon in Chapter 6.7 and 6.8.

6.4. Differential Bias Correction

Intensity bias or intensity inhomogeneity is the slowly changing and smooth spatial variation that can occur within any MRI scan. Grey and white matter in one part of the brain can have systematically lower signal intensity than at another part of the scan. This inhomogeneity causes significant problems with image analysis and consistency in automated segmentation and image registration.

The artefact is caused by (1) inhomogeneity of the magnetic field (B_0) of the MR system. (2) inhomogeneity of the radio-frequency (RF) pulse generated by the oscillating secondary magnetic field (B_1). (3) non-uniform sensitivity of the receiver coils used to detect the MR signal.

Automated segmentation techniques require a model for signal intensity for different tissue types. These models become invalid for images with a large bias. Manual segmentation, using thresholding may also be affected.

Pre and post processing may be used to correct bias. Pre processing techniques require increased acquisition time and cannot be used to correct retrospective data.

Linear registration which relies on the brain boundary shift integral (BBSI) will be affected by intensity changes and deformation based atrophy techniques will also be affected as the artificial intensity change will produce unrealistic warps. The BBSI will not give a valid measure of atrophy when there is significant tissue inhomogeneity between the scans. Hence, for all data analysis in this thesis, MRI scans were corrected for inhomogeneity artefact using differential bias correction (DBC) (Lewis and Fox, 2004).

In order to determine whether differential bias correction improved the repeatability and reliability of regional segmentation, segmentation of the third ventricle was undertaken on non-DBC scans and on DBC scans, using the same segmentation

protocol. The third ventricle was chosen as it is a CSF space and therefore rapidly segmented using semi-automated threshold based measurement techniques, with minimal operator input. This should minimise measurement error and allow a more accurate assessment of whether DBC influences the accuracy of the segmentation.

In describing the extent to which the instruments agree with each other, three issues are of importance:

1) Bias: On average, are measurements made with the second method, higher or lower than those made with the first? If so, what is the extent of the bias?
2) Differences in variance: Are measurements made with the second instrument more variable than those made with the first? This method concerns the similarity between within subject standard deviation (σ_{w1}^2 and σ_{w2}^2).
3) Variability of differences: How variable are the differences between measurements made on the same individual using different instruments? This question concerns the size of σ_{w1}^2 and σ_{w2}^2.

6.4.1. Methods

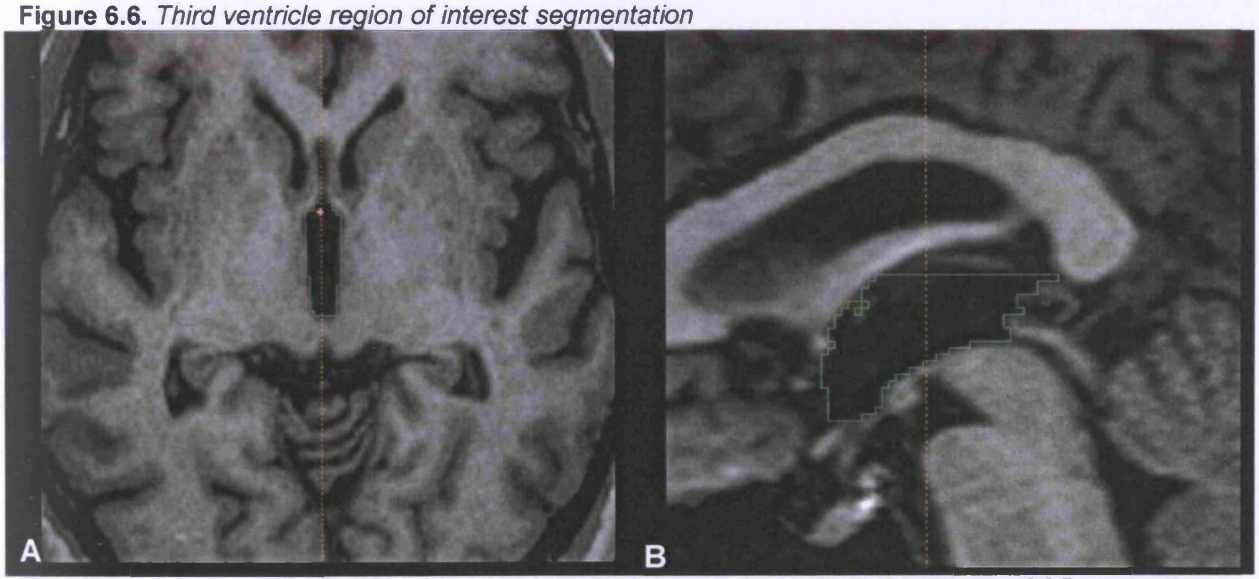
For third ventricle measurements, all of the above points were addressed with respect to volume measurements using MRI scans registered to standard space: (1) with; and (2) without differential bias correction (DBC) to determine whether this post-processing tool improved repeatability and reliability of the regional volume measurements.

The third ventricle volume was measured by a single rater blinded to the identity of the individual, the clinical diagnosis and the time sequence of the scan, in 50 subjects randomly selected from the data set. To minimise practice effect, the third ventricle was segmented on DBC and non-DBC MRI scans with a six-week interval.

Each subject had the third ventricle volume segmented twice, for the scan at time point one and time point two and for non-DBC and DBC scans (hence 8 segmentations).

For ventricular segmentation, the objective is to exclude brain voxels. Using a threshold based automated segmentation proceedure (Freeborough *et al.,* 1997), the upper threshold was set at 60% of mean brain intensity and the lower threshold at zero in order to exclude voxels that were predominantly brain tissue. The anterior border was set at the anterior commisure and the posterior border at the anterior border of the splenium of the corpus callosum in the mid sagittal slice. The inferior border was set at the level of the mamillary bodies. Voxels within the aqueduct of sylvius were excluded. A horizontal line at the level of the inferior border of the splenium of the corpus callosum, in the mid sagittal slice was set as the superior border of the third ventricle. These definitive cut-offs were set in order to increase the repeatability and reliability of the volume measurements. Examples of third ventricle segmentations are shown in figure 6.6A and B.

Figure 6.6. *Third ventricle region of interest segmentation*



T1 weighted MRI in MIDAS window showing seed (pink cross in A) planted within the CSF space. A- Axial third ventricle ROI. B- Sagittal third ventricle ROI. The anterior cut off at the anterior commissure is clearly demonstrated as is the posterior and superior border defined by the splenium of the corpus callosum.

For all subjects with a complete data set, the within subject standard deviation (σ_w) was calculated as a measure of repeatability for the non-DBC and the DBC scans at time point one and two. A Pitman's test was performed to test the hypothesis that the non-DBC and the DBC measurements have the same variance (null hypothesis $\sigma_{w1}^2 = \sigma_{w2}^2$).

The reliability of the methods used was evaluated by calculating the between person variance in each measurement method and from this, the reliability coefficient (or ICC) was calculated.

Bias in the measurement techniques was assessed by using a one-sample t-test applied to differences in the mean measurements at each time point, comparing the non-DBC segmentations with the DBC segmentations, with a null hypothesis that the true mean difference was zero.

6.4.2. Results

Of the 50 subjects studied, eight sets of segmentation data were included for evaluation of third ventricle measurements.

(1)Third ventricle, time 1, segmentation 1.(5)DBC third ventricle, time 1 segmentation 1.

(2)Third ventricle, time 1, segmentation 2.(6)DBC third ventricle, time 1 segmentation 2.

(3)Third ventricle, time 2, segmentation 1.(7)DBC third ventricle, time 2 segmentation 1.

(4)Third ventricle, time 2, segmentation 2.(8)DBC third ventricle, time 2 segmentation 2.

Table 6.3 *Third ventricle segmentation data*

	n	Mean (ml)	SD	Min	Max
3v11	50	1.95	.81	.63	3.95
3v12	50	1.99	.83	.63	4.07
3v21	50	2.03	.85	.64	4.10
3v22	50	2.07	.86	.70	4.20
DBC3v11	50	2.00	.83	.63	4.04
DBC3v12	50	2.00	.82	.64	3.94
DBC3v21	50	2.08	.87	.69	4.16
DBC3v22	50	2.07	.86	.69	4.17

3v11-Third ventricle, time 1, segmentation 1 ; DBC3v11-DBC third ventricle, time 1 segmentation 1.
3v12-Third ventricle, time 1, segmentation 2 ; DBC3v12-DBC third ventricle, time 1 segmentation 2.
3v21-Third ventricle, time 2, segmentation 1 ; DBC3v21-DBC third ventricle, time 2 segmentation 1.
3v22-Third ventricle, time 2, segmentation 2 ; DBC3v22-DBC third ventricle, time 2 segmentation 2.

A simple model of repeatability implies that for two measurements per subject, the observed group differences should have zero mean and constant small variance.

The within subject standard deviations, and the results of the variance ratio test are shown in table 6.4.

Table 6.4 *Repeatability data for third ventricle measurements*

	Within subject		***Variance ratio test***		
	SD	**Mean difference (ml) (95%C.I)**	***Null Hypotheses***		
			3v1 ≠ DBC 3v1	***3v1 < DBC 3v1***	***3v1 > DBC 3v1***
3v1	0.061	0.06 (0.04-0.08)	P=0.0014	P=0.99	P=0.0007
DBC3v1	0.038	0.04 (0.03-0.05)			
			3v2 ≠ DBC 3v2	**3v2 < DBC 3v2**	**3v2 > DBC 3v2**
3v2	0.065	0.07 (0.05-0.09)	P=0.0054	P=0.99	P=0.0027
DBC3v2	0.038	0.04 (0.03-0.05)			

3v1- non DBC third ventricle measurements at time point 1; 3v2- non DBC third ventricle measurements at time point 1; DBC3v1- DBC third ventricle measurements at time point 1; DBC3v2- DBC third ventricle measurements at time point 2; sd-standard deviation; Mean difference- mean difference between segmentation 1 and 2.

The reliability of the method for measuring the volume of the structures can be calculated from a one-way analysis of variance between the first and second measurements of the third ventricle volume at each time point for the non-DBC and

the DBC scans. This allows calculation of the between subject variance and the within subject variance, and enables the ICC to be derived.

Table 6.5 *Reliability coefficients for third ventricle measurements*

	Reliability coefficient (ICC)
3v1	0.994
3v2	0.994
DBC 3v1	0.998
DBC 3v2	0.998

From this data it is apparent that both methods are reliable (ICC close to 1). The DBC scan segmentation has a slightly greater ICC.

Analysis of the mean of the 3v1, DBC3v1, 3v2 and the DBC3v2 measurements and the differences between measurements of the non-DBC and DBC scans at the same time point suggests that the non-DBC measurement method has a tendency (a bias) to give a lower volume for the third ventricle using the same measurement protocol in every other respect. The results of a one-sample t-test are shown in the table below, confirming this suspicion.

Table 6.6 *Statistical comparison of methods of third ventricle measurement*

	Volume (ml) (SD)	Difference in means (ml)	95% C.I	Null hypothesis	
				Difference=0	Difference>0
Mean 3v1	1.97 (0.82)	0.027	0.006-0.05	P=0.01	P=0.005
Mean DBC3v1	2.00 (0.83)				
Mean 3v2	2.05 (0.86)	0.024	0.008-0.04	P=0.004	P=0.002
Mean DBC3v2	2.07 (0.87)				

6.4.3. Conclusions

These data confirm that the methods of third ventricle segmentation have excellent intra-rater repeatability and reliability. The results of a Pitman's test on third ventricle measurements suggest that the repeatability and reliability are improved when using scans that have been corrected for differential bias. Segmenting the non-DBC scans

has a bias to returning a lower volume for the third ventricle compared to segmentation of the DBC scans.

The improved repeatability and reliability for segmenting this region with DBC is difficult to explain on the basis of this post acquisition processing technique alone. The difference in the ICC is minimal suggesting the reliability of the measurement is only minimally influenced by DBC.

DBC should reduce artefactual intensity differences and as such, thresholding may be affected, particularly around brain and CSF boundaries. This could help to explain the improvement in repeatability and lead to the bias towards a larger volume recorded for the DBC scan segmentations.

The main difficulty is in discounting the intrarater practice effect as a causal factor in these improvements. A six-week gap was left between segmenting the regions in the non-DBC and the DBC scans, although despite this, some practice effect is likely to remain. As the DBC measurements were conducted after the non-DBC scan measurements, it could explain the reduced within subject standard deviation but this would not contribute to the bias which is also seen.

Fully automated regional segmentation is possible but whilst this results in excellent repeatability and reliability, the accuracy of regional measurement is not as good as manual, semi-automated techniques. For this reason manual, semi-automated segmentation was used for ROI studies in this thesis.

DBC scan pairs registered to standard space were used for all further whole brain and regional volume segmentations, to ensure the best repeatability and reliability.

6.5. Whole brain and region of interest (ROI) segmentation

6.5.1. Methods

All scans were registered to standard space and were segmented as scan pairs with baseline and repeat scans viewed together with the operator blinded to the identity and diagnosis of the individual as well as the temporal sequence of the images being viewed. The results section in this chapter (6.5.4) only describes cross sectional volume analysis.

ROI segmentation relies partly on manual application of boundaries. For all scans a semi-automated intensity-threshold based processing routine was used to delineate regions that were "brain" and "non-brain". This involves using the MIDAS software tool (Freeborough *et al.,* 1997) to "plant" a seed within the ROI on each scan slice. In this way, the brain/CSF boundary can be delineated without manual outlining. Each scan slice is reviewed and manually adjusted to include only the desired region on the basis of anatomical knowledge of the regional boundaries.

Individual MR images were segmented with the operator blinded to the identity and diagnosis. Specific methodology for whole brain and regional segmentation is detailed in the following sections.

6.5.1.1.Brain volume

The aim of whole brain segmentation is to label the 3D brain while excluding non-brain (i.e. scalp, CSF, dura and the superior sagittal sinus). The following four steps describe the segmentation process in detail (Freeborough *et al.,* 1997).

<u>Step 1</u> - Setting the threshold values

1. An upper threshold value was set and any voxels brighter than this upper threshold excluded. The lower threshold is set at zero (figure 6.7a). Susceptibility

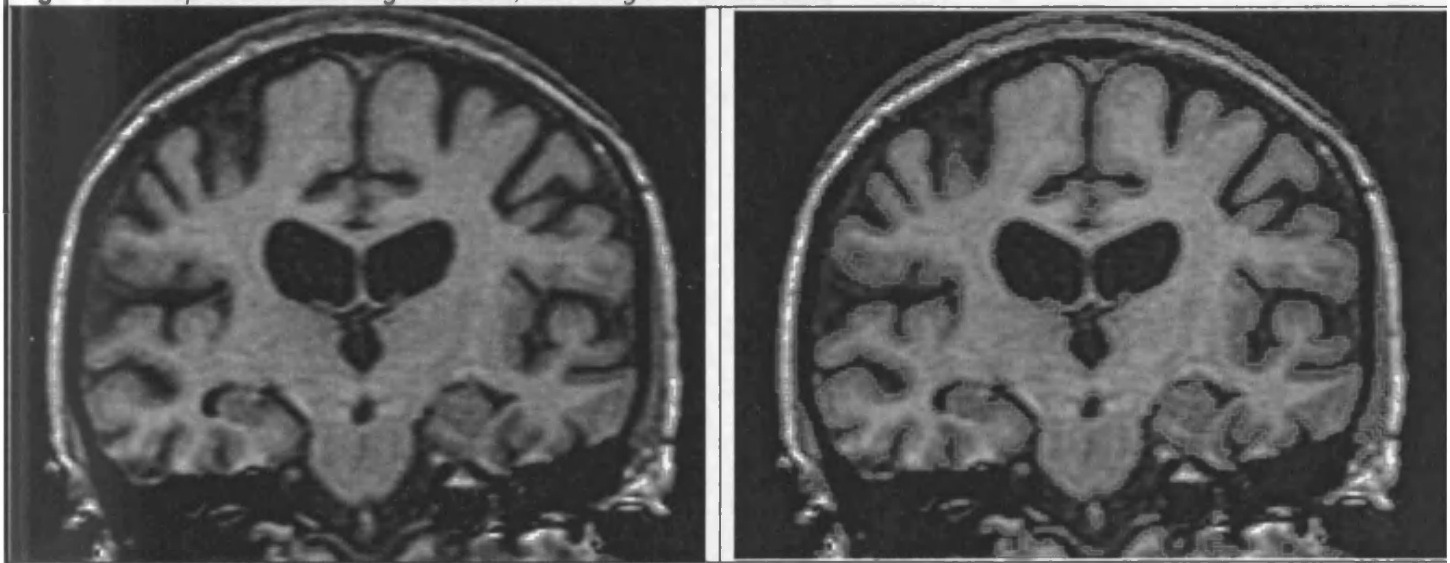
artefact occasionally appears around the fusiform and inferior temporal gyri in the temporal lobe as abnormally bright tissue. Any susceptibility artefact is excluded but normal white matter remains.

2. The lower threshold is then set to just over half the value of the upper threshold. Anything darker than the lower threshold will be excluded.

   The aim is to outline the brain as accurately as possible. The green intensity region should now include most of the brain. The brain is checked to ensure that the cerebellum and the temporal lobes are included. Ideally the number of bridges (connections between brain and dura which lie within the intensity range defined for brain) should be zero. Inevitably a compromise must be struck but, ideally, the lower threshold should be taken to a high enough value so that in the next step only one erosion is needed to remove non-brain and leave only brain (figure 6.7b).

3. The final part of this step is to set the axial cut off slice. This removes any non-brain below the cerebellum and aims to give a reproducible lower/inferior limit to what is included in the brain stem.

Figure 6.7 *Step one in brain segmentation, removing non-brain tissue*



A- exclusions of voxels above the upper threshold

B- outlining of brain voxels with minimal bridges between brain and non-brain tissue

Step 2 - **Erosions**

This step eats away at the external surfaces of the brain region that has been outlined above until only brain remains (figure 6.8a). The upper threshold value set in step 1 is used to limit the erosion to tissue with a signal intensity below the threshold.

Figure 6.8a *Removal of non-brain tissue*

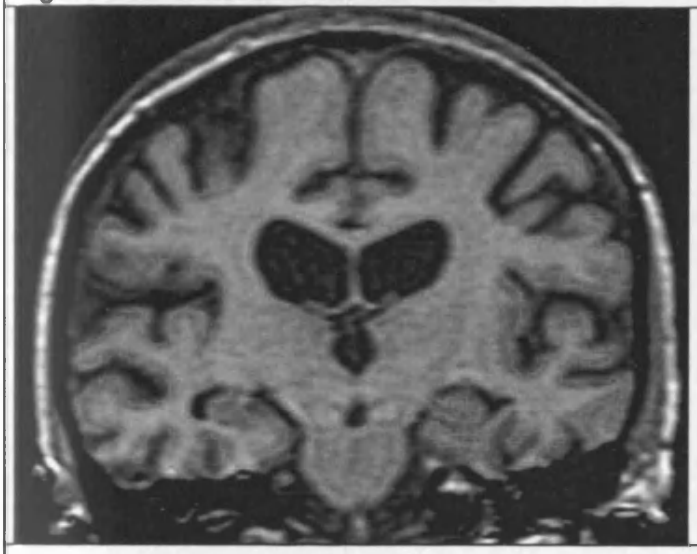
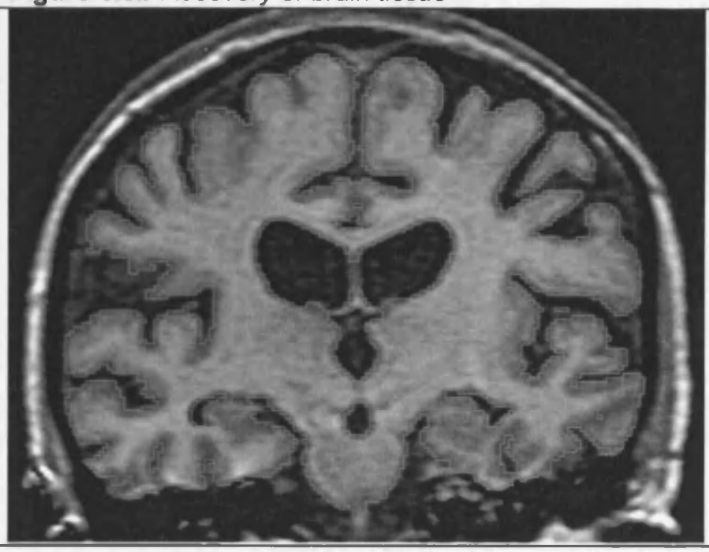


Figure 6.8b *Recovery of brain tissue*



Step 3 - **Dilations**

The purpose of this step is to recover the brain removed by the erosion process. Each dilation reclaims pixels that fall within the permitted threshold range and are adjacent to the brain region (figure 6.8b). The thresholds that appear at this point are automatically set by calculating the mean signal intensity over the whole brain region after step 2, taking 60% and 160% of this value. These thresholds are re-calculated after each dilation as more voxels are recovered.

Step 4 – **Re-thresholding box**

This step completes the segmentation process. The re-thresholding box is an invisible box of voxels that moves across the outlined brain region filling in any area within that region that has been omitted by step 3. However, there are three rules which govern this procedure. The first is that the omitted area must be smaller than the box.

Secondly, only those voxels within this area that have a signal intensity of between 60 and 160% of the mean will be included. Finally, the missing area must already be included within the region. Therefore, any voxels outside the brain region will not be recovered. Generally, a rethresholding box size of 6 voxels is large enough to recover any areas previously excluded.

Finally, axial and sagittal views are checked one final time to ensure accuracy of the regional segmentation.

6.5.1.2.Midbrain

For midbrain segmentation, two orthogonal views were available. A consistent lower threshold of 70% of mean brain intensity was applied to exclude voxels with a lower intensity that were predominantly cerebrospinal fluid (CSF). The superior border of the midbrain was taken as the upper border of the midbrain tegmentum in the mid-sagittal slice and the inferior cut-off taken at the level of the superior border of the pons in the mid-sagittal slice. The posterior and anterior borders were defined by the brain tissue/CSF boundary. Higher up in the midbrain, where no CSF border was present, a manual outline of the grey white tissue boundary was made. Returning to sagittal slices, care was taken to ensure that voxels within the quadrigeminal plate were included. Examples of midbrain volume segmentation are displayed in figure 6.9A and B.

6.5.1.3.Pons

As for other regions, the segmented brain was registered into standard space. The lower threshold was set at 70% of the mean brain intensity, and the upper threshold set at maximum intensity. The mid sagittal slice was selected in the main MIDAS window. A seed was inserted in the pons and the anterior convex border and the posterior border delineated by thresholding. The upper border was set at the level of a

horizontal line extending from the posterior upper-most tip of the pons convexity. The lower border was set at the level of a horizontal line extending from the posterior and lower most tip of the pons convexity. The uppermost axial slice was then selected and a horizontal line drawn through the median suclus on the posterior aspect of the brain stem. This was taken as the posterior border of the pons in the axial view.

When the middle cerebellar peduncles become apparent, the lateral border is taken as a vertical line medial to the emergence of the trigeminal nerve. Examples of segmentation of the pons are shown in figures 6.9E and F.

6.5.1.4.Cerebellum

For the cerebellum, a lower threshold at 60 % of the mean brain intensity was set with the upper threshold set at maximum. The cerebellum was segmented from anterior to posterior. Voxels within the cerebellar peduncles (superior, middle and inferior) were only included if, on the axial slice, cerebellar tissue was visible in the midline, connecting the cerebellar hemispheres. Examples of cerebellar segmentation are shown in figure 6.9C and D.

6.5.1.5.Superior cerebellar peduncle

The SCP was segmented using the protocol outlined earlier in this chapter (section 6.3 and figure 6.2)

6.5.1.6.Lateral ventricle

The protocol for lateral ventricle segmentation is well established (Dalton *et al.*, 2002). The aim of the segmentation is to label all voxels of ventricular cerebrospinal fluid, but not extraneous CSF spaces or brain tissue. The coronal view was selected in the main MIDAS window. The upper threshold was set at 60% of the mean brain voxel intensity and the lower threshold at zero.

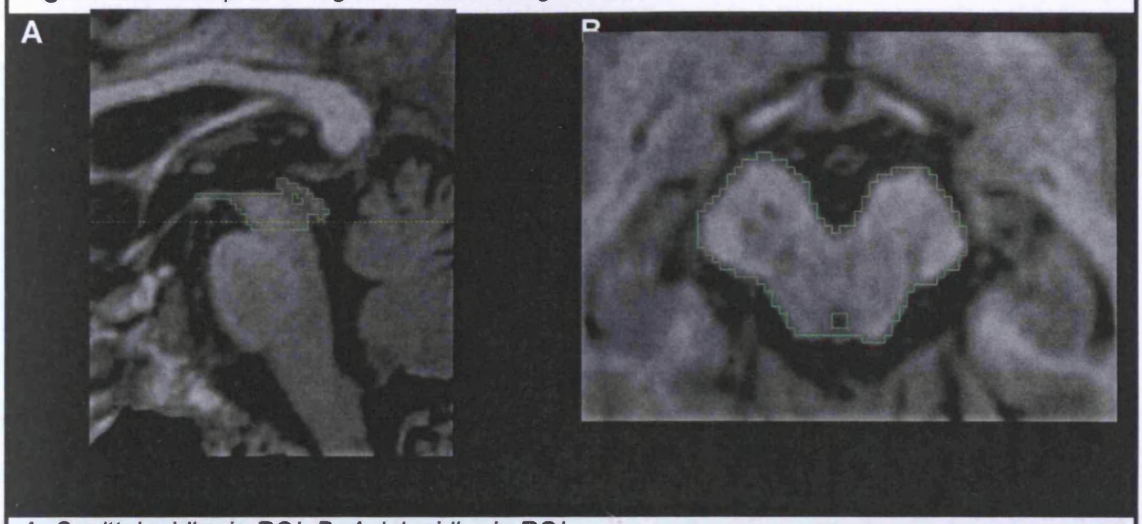
All the ventricle regions are included and the temporal horn specifically checked to ensure its inclusion.

Each coronal slice was viewed, adding the lateral ventricle regions moving anteriorly through the brain. Where the region spills out into surrounding CSF spaces, the border of the ventricle was manually corrected. The third ventricle was manually excluded. Finally all slices were checked to ensure only voxels within the lateral ventricles are included. Examples of lateral ventricle segmentation can be seen in figure 6.10.

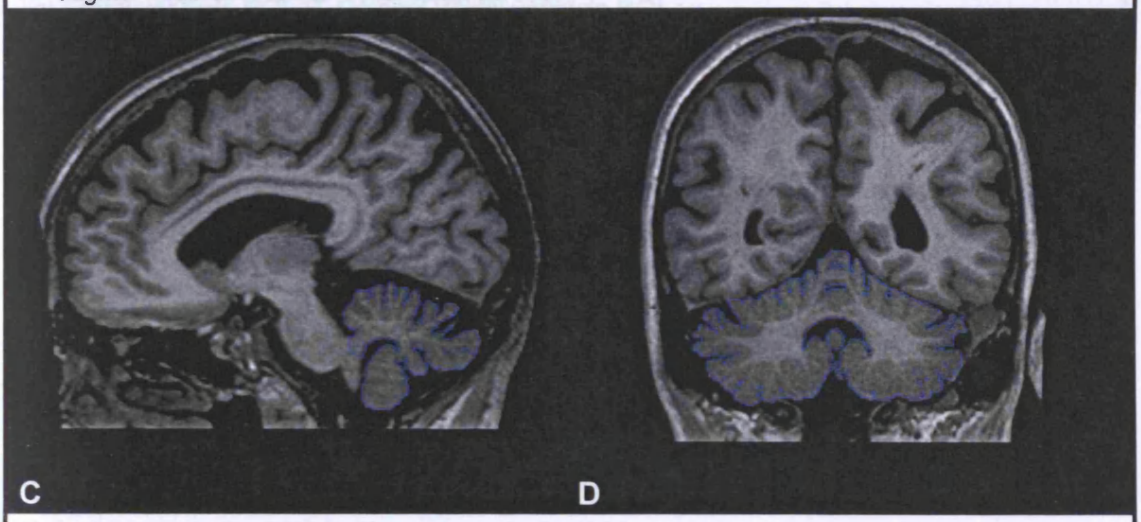
6.5.1.7.3rd Ventricle

Segmentation of the third ventricle was undertaken using the protocol described in section 6.4 (figure 6.6).

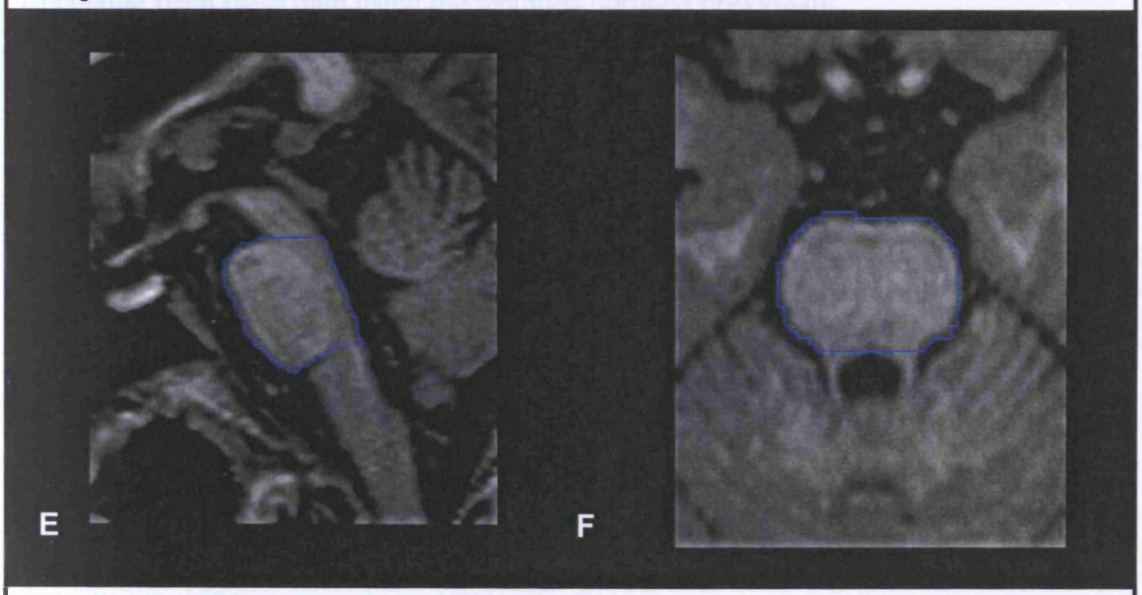
Figure 6.9. *Examples of region of interest segmentation*



A- Sagittal midbrain ROI. B- Axial midbrain ROI

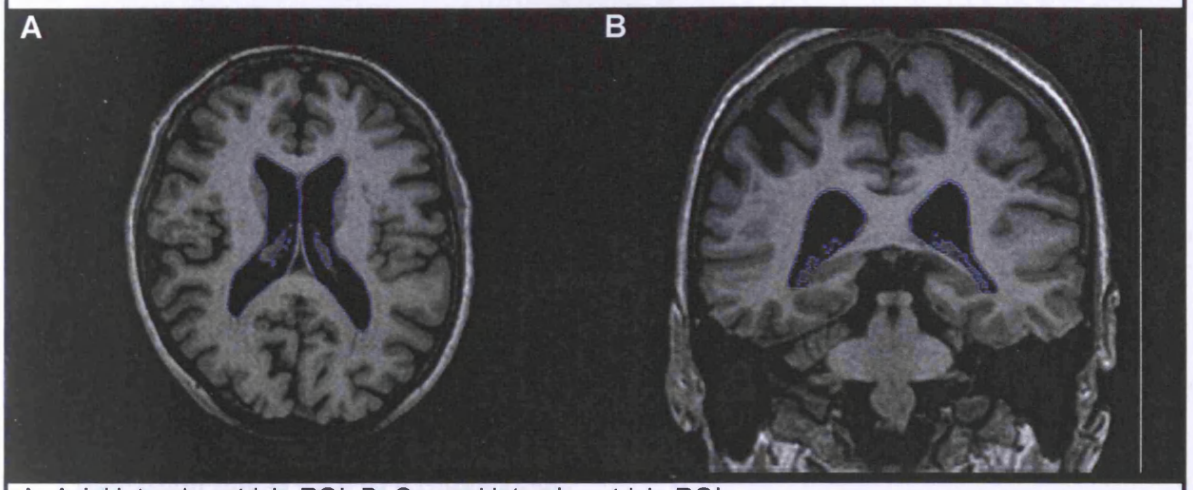


C-Sagittal cerebellar ROI. D- Coronal cerebellar ROI.



E-Sagittal Pons ROI. F- Axial Pons ROI.

Figure 6.10. *Ventricular Region of Interest Segmentation*



A- Axial lateral ventricle ROI. B- Coronal lateral ventricle ROI

6.5.2. Repeatability and reliability of regional segmentation

Repeatability and reliability of whole brain and lateral ventricle measurements are well established (Freeborough *et al.,* 1997; Whitwell *et al.,* 2001). For midbrain, pons and cerebellar volume measurements, after initially segmenting the volumes on the entire cohort of subjects, 10 MRI scans were reassessed using the same methods, with the rater blinded to the diagnosis. Repeatability and reliability measures were calculated from these data using the methods outlined previously.

Repeatability and reliability analysis of midbrain measurements, with both sets of blinded measurements conducted using DBC scan pairs are reported in table 6.7.

For assessment of repeatability and reliability scores relating to midbrain, pons and cerebellar volume measurements, four sets of data were evaluated.

(1)Region time1, segmentation 1.

(2)Region time1, segmentation 2.

(3)Region time2, segmentation 1.

(4)Region time2, segmentation 2.

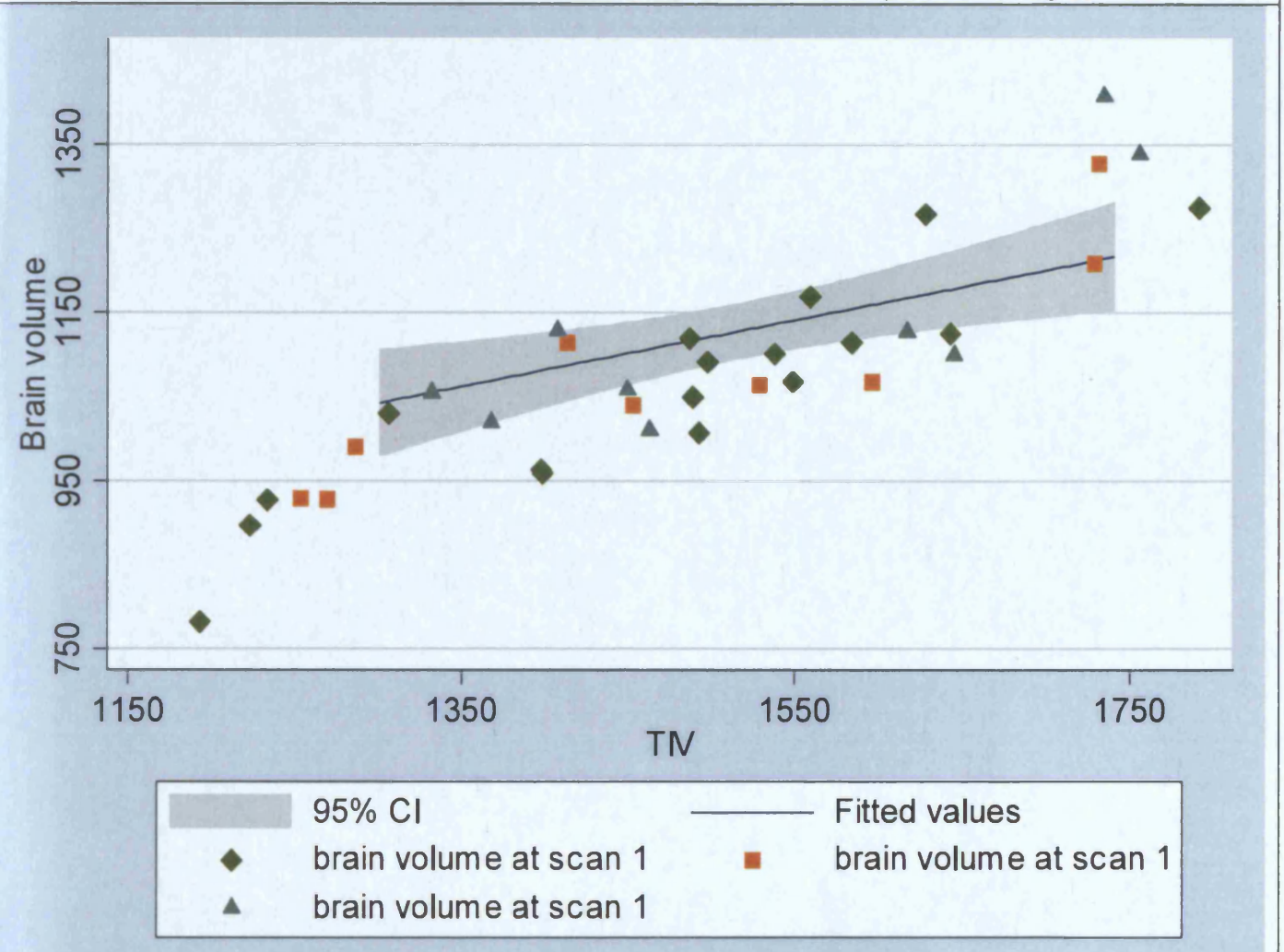
Hence, for all structures, if the individuals TIV was larger than the mean TIV in controls, correction results in a reduced volume of that structure.

Normalizing the brain volumes in this way reduces the scatter of data and reduces sex-related differences in brain volumes (Whitwell *et al.*, 2001).

6.5.3.3.Results

Figure 6.11 demonstrates that the disease groups have a greater degree of brain atrophy compared to healthy controls when the healthy control brain volume-TIV regression line is used as a rough guide.

Figure 6.11 *Relationship between brain volume and TIV in PSP, MSA and PD compared to healthy controls.*



Y axis-Brain volume (ml), X axis-TIV-total intracranial volume (ml). Green (diamond) - PSP subjects, Orange (square) – MSA-P, Grey (triangle) -PD. Line and 95% confidence interval represents the relationship between Brain volume and TIV in healthy controls.

In the group as a whole, the mean TIV in men was 11% greater than in women ($p<0.0001$), in keeping with previous studies (table 6.9).

Table 6.9 *Age and TIV in all subjects and in controls*

	n (male:female)		Mean age (SD) (years)	meanTIV ± sd (ml)	95% C.I
Whole group	55		65.1 (6.4)	1504 (149)	1464-1544
	♂	32	65.3 (6.0)	1577 (125)*	1532-1622
	♀	23	64.9 (7.1)	1403 (117)	1352-1454
Healthy Controls	18		66.7 (5.4)	1514 (125)	1452-1577
	♂	10	67.6 (3.9)	1586 (111)**	1507-1666
	♀	8	65.7 (6.9)	1424 (74)	1363-1486

* p<0.0001♂>♀; ** p=0.003 ♂>♀

Table 6.10 *Age and TIV in healthy controls and disease subsets*

	n (male:female)	Mean age (SD)(years)	meanTIV (SD) (ml)	95% C.I
Healthy controls	18 (10:8)	66.7 (5.4)	1514 (125)	1452-1577
PSP	18 (11:7)	65.1 (5.7)	1489 (158)	1410-1566
MSA-P	9 (6:3)	62.3 (8.0)	1484 (183)	1345-1625
PD	9 (5:4)	64.3 (8.3)	1545 (159)	1423-1667

There were no significant differences in the TIVs between healthy controls, PSP, MSA or PD subjects (One way analysis of variance, F=0.36, p=0.7, table 6.10).

Plotting the logs of the baseline brain and regional volumes against the log of the TIV in controls established the slope of the relationship between TIV and the volume of each structure. The resulting coefficient (table 6.11) was used to correct the volume of each structure.

Table 6.11 *List of coefficients from regressing ln(volume) and ln(TIV) in control subjects.*

Structure		Grad
Midbrain	1	-0.05
	2	-0.05
Pons	1	0.42
	2	0.43
Superior cerebellar peduncle	1	0.17
	2	0.18

ln-natural logarithm, 1-first scan, 2 follow up scan.
Grad- regression coefficient (slope) of ln(structure) and ln(TIV)

6.5.3.4.Conclusions

Total intracranial volume measurements in controls were consistent with those published in previous studies (Whitwell *et al.,* 2001). These data confirm the well-recognized wide range in intracranial volumes between individuals. The smallest volume was less than 70% of the largest TIV. On average, the men had TIVs that were 11% larger than those of the women. In conclusion, these adjustments are important in studies assessing group differences in total and regional cerebral volume, which otherwise may be confounded by differences in head (and TIV) size between groups. The coefficients from regressing ln(volume) and ln(TIV) in the control subjects are the same or very close to the same value for baseline and follow up scans. However, while the value for the midbrain and the SCP is close to zero, suggesting little relationship with a changing TIV, that for the pons is larger suggesting it increases with an increasing TIV.

6.5.4. Results of Region of interest segmentations

For normally distributed data, a one-way analysis of variance, and for non-parametric data a Mann-Whitney U test, were used to determine group differences in brain, third ventricle and lateral ventricle volume as well as, midbrain, pons, cerebellar and SCP volume, all corrected for TIV as described above. There were significant differences in frontal, midbrain, SCP and third ventricle volumes between PSP and MSA, PD and healthy controls. Significant differences were seen in pons and cerebellar volumes between MSA and PSP, PD and controls (tables 6.12 and 6.13 and figure 6.12).

For all regions where the mean volume differentiated one disease group from another or from controls, forward stepwise logistic regression using a likelihood ratio test (with $p<0.05$ as criteria for addition to the model), was performed to determine

whether these volumes were useful predictors of the clinical diagnosis when considered together. The results of this analysis are reported in section 6.7.

Table 6.12. *ROI segmentation mean (SD) volumes (TIV uncorrected)*

	Subject Group ROI volumes (ml) (mean ± sd)			
	PSP	**MSA-P**	**PD**	**Control**
Whole brain volume	1052.4 (124)	1074.2 (130)	1138.2 (141)	1128.9 (82.4)
Cerebellar volume	109.6 (17.3)•	**89.9 (27.3)◄**	126.1 (14.3)	119.8 (13.5)
Midbrain volume	**5.79 (1.0)◄/****	**7.08 (1.2)▲**	7.87 (1.1)	8.33 (0.8)
Pons volume	12.8 (1.7)▼	**9.60 (4.3)#**	15.4 (2.6)	13.9 (1.5)
SCP volume	**0.41 (0.08)►**	0.54 (0.05)	0.55 (0.05)	0.55 (0.06)
LV volume	38.5 (16.1)	33.7 (17.0)	36.2 (18.1)	26.8 (13.1)
Third ventricle volume	**2.5 (0.8)◊**	1.82 (0.6)	1.96 (0.9)	1.59 (0.7)

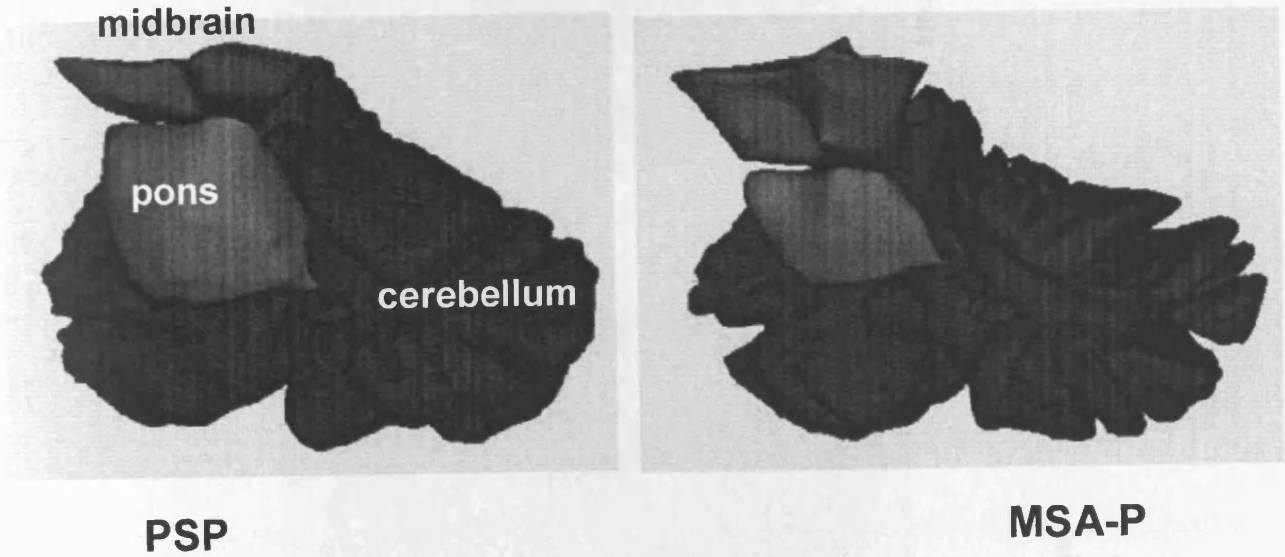
Values with significant differences are shown in bold. •-p=0.02 vs. PD; ◄-p=0.05 vs. PSP, p=0.003 vs. PD, p=0.005 vs. controls; *-p=0.007 vs. MSA, ** p<0.001 vs. PD and controls; ▲-p=0.004 vs. controls; ▼-p=0.01 vs.PD; # -p=0.02 vs. PD/controls; ►-p<0.001 vs. MSA-P/PD/controls; ◊-p=0.002 vs. control.

Table 6.13. *ROI segmentation mean (SD) volumes (TIV corrected)*

	Subject Group ROI volumes (ml) (mean ± sd)			
	PSP	**MSA-P**	**PD**	**Control**
Whole brain volume	1085.9 (57.2)	1109.8 (59.9)	1130.4 (78.7)	1132.5 (74.3)
Cerebellar volume	110.2 (15.9)•	**90.2 (26.9)****	125.9 (13.1)	119.9 (13.4)
Midbrain volume	**5.69 (1.1)◄/****	**7.08 (1.2)▲**	7.87 (1.1)	8.33 (0.8)
Pons volume	12.8 (1.6)	**9.60 (4.3)****	15.4 (2.5)	13.9 (1.5)
SCP volume	**0.40 (0.08)****	0.54 (0.05)	0.55 (0.05)	0.55 (0.06)
LV volume	39.3 (14.9)	34.4 (15.3)	36.0 (18.0)	26.7 (12.3)
Third ventricle volume	**2.52 (0.8)◊**	1.93 (0.6)	1.92 (0.8)	1.59 (0.6)

Values with significant differences are shown in bold. •-p=0.04 vs. PD *-p=0.009 vs. MSA, **-p<0.001 vs. PD/Control (and MSA-P for SCP volume), (and p=0.01 vs. PSP for cerebellar volume), ▲-p=0.03 vs. controls; ◊-p=0.001 vs. controls.

Figure 6.12 *Segmented brainstem and cerebellar volumes in PSP and MSA-P demonstrating relative differences in midbrain, pontine and cerebellar volume, within and between the diseases.*



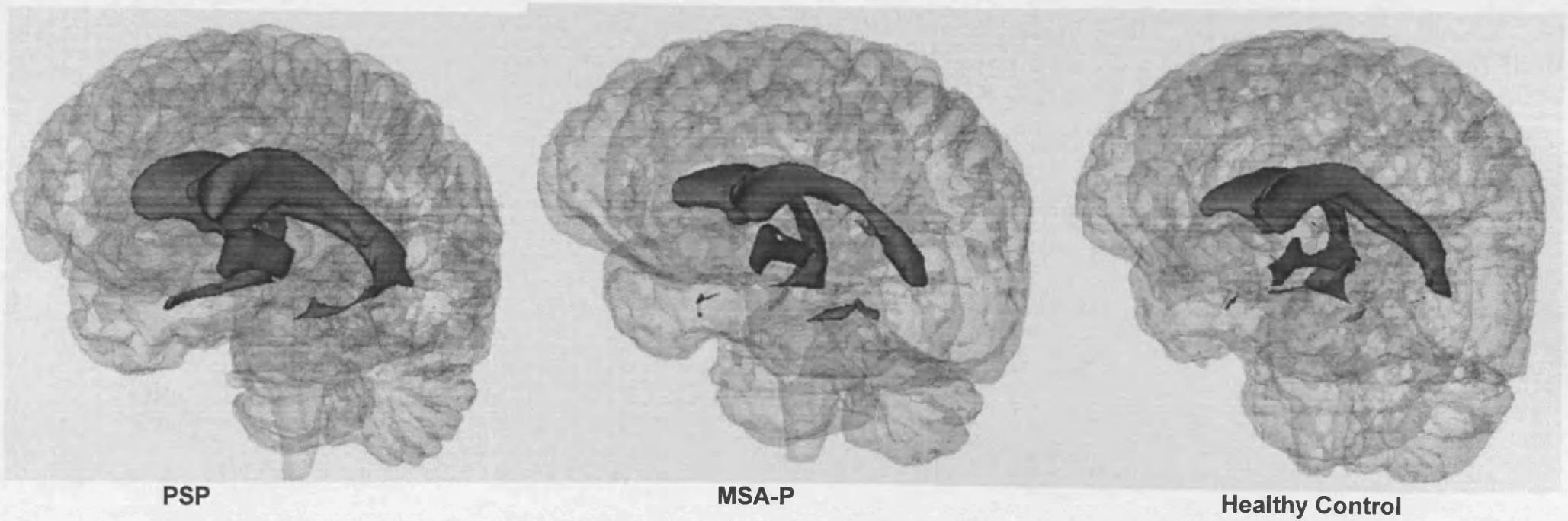


Figure 6.13 *Segmentation of lateral ventricles and third ventricle superimposed on brain volume.*

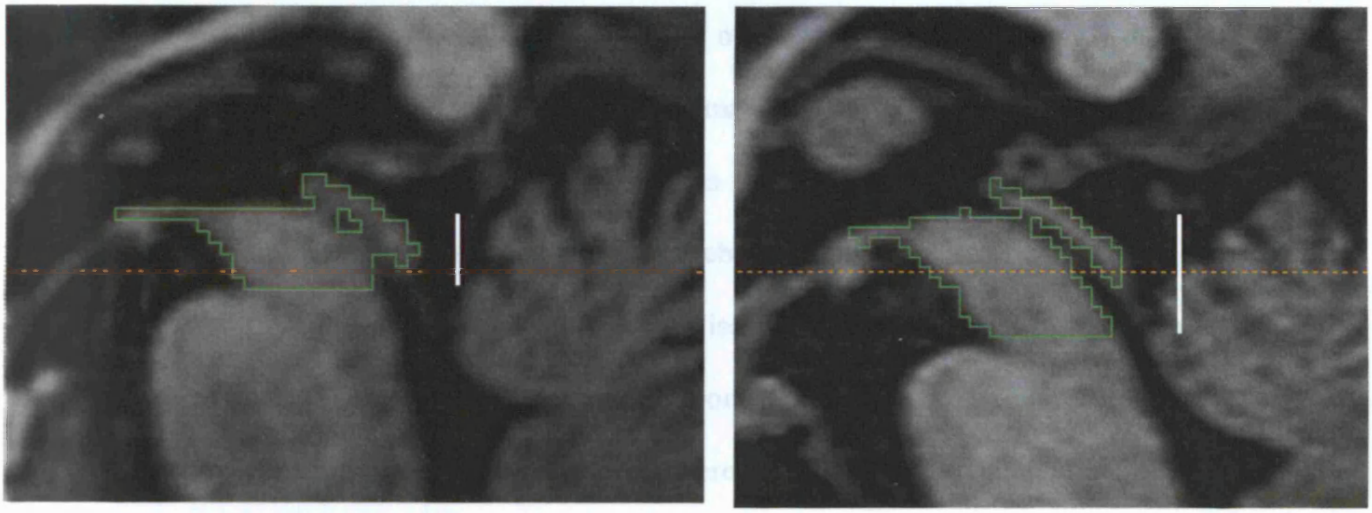
6.5.5. Conclusions

In this prospective MRI based study comparing volumetric ROI in PSP, MSA-P, PD and healthy controls, reliable and reproducible manual and semi-automated ROI volume analysis confirm that PSP and MSA-P have differences in regional brain volume which distinguish them from PD and healthy controls and more importantly from each other. Correcting the volumes for TIV has little effect in terms of the difference in the mean values between the groups. Midbrain volume in PSP is approximately 80% of the midbrain volume in MSA-P and ponto-cerebellar volume in MSA-P is approximately 80% of the ponto-cerebellar volume in PSP. It is only the larger SD in pontine volume in MSA-P (possibly due to measurement error) that results in a less significant difference between MSA and PSP.

The results from this study support previous retrospective and qualitative studies suggesting that midbrain size is significantly reduced in PSP. Interestingly it is also reduced in MSA-P, perhaps reflecting the known pathological involvement of the subtantia nigra and subthalamic nucleus in this disease.

Comparing the shape of the midbrain on MRI in the PSP subjects with healthy controls, suggests that it is the height of the midbrain tegmentum that is lost in PSP resulting in a reduction in volume (figure 6.12 and 6.14). This macroscopic atrophy is also frequently commented on in post mortem reports of PSP cases at the Queen Square Brain Bank. The reduction in height possibly occurs as a consequence of pathological involvement and subsequent tissue loss in the subthalamic nucleus, substantia nigra and red nucleus.

Fig 6.14 *Reduction in Midbrain tegmentum height in PSP. The vertical white line demonstrates height.*



PSP Healthy control

Reduced pontine and cerebellar volumes distinguish MSA-P from PSP, PD and healthy controls. That a reduced cerebellar volume is observed in MSA-P is not a surprise as this is a finding often-cited in previous qualitative imaging studies. However, our cohort of MSA patients includes only those with a predominant parkinsonian phenotype at entry into the study and despite this, a significantly lower mean cerebellar volume, compared to healthy controls, PD and PSP was detected. This suggests that significant cerebellar and pontine pathology resulting in atrophy is present in MSA even when a Parkinson's syndrome is the presenting phenotype. This is supported by the results of a recent pathological study in which the severity of pontine pathology did not differentiate predominantly cerebellar (MSA-C) from the predominantly Parkinsonian forms (MSA-P) (Ozawa *et al.*, 2004).

Cerebellar volumes in PSP were lower than in PD. Cerebellar cortical pathology is minimal in PSP although the dentate nucleus of the cerebellum is lesioned. However atrophy of this structure would not be expected to result in a large volume of tissue loss as relative to the rest of the cerebellum its volume is small. There is an absence of

cerebellar signs in PSP, and in fact their presence would exclude the diagnosis on clinical grounds. However as the cerebellar outflow tracts are consistently damaged, this lack of clinical signs is surprising. It may be that as in MSA, the presence of significant pathology in other brain regions such as the basal ganglia, resulting in extrapyramidal motor deficit, masks any cerebellar signs.

Despite the fact that brain volumes in the diseased subjects seemed to fall below the expected confidence interval for healthy control brain volume-TIV regression line (figure 6.9), there was not a significant difference between the whole brain volume in PSP, MSA-P, PD or healthy controls, consistent with relatively localised subcortical loss rather than very generalised atrophy (see also Chapter 6.6).

In contrast to an earlier study from the National Hospital for Neurology and Neurosurgery (Schrag *et al.,* 2000), the third ventricle volume only differentiated PSP from healthy controls and not from MSA-P or PD. This suggests that an enlarged third ventricle should not be considered as supporting a diagnosis of PSP.

Dilatation of the lateral ventricles is unlikely to result simply as a consequence of tissue loss from peri-ventriclar regions. More probable is that it occurs as a consequence of tissue loss from the frontal cortex and white matter with smaller contributions from the temporal lobes. There is no statistically significant difference between lateral ventricle volumes in PSP, MSA-P, PD or healthy controls. This is despite a 40% larger lateral ventricle volume in PSP than in healthy controls. The wide natural variability in ventricular volume probably accounts for this lack of a significant difference.

6.6. Segmentation of frontal quadrants and posterior-inferior regions

6.6.1. Introduction

Accurate volumetric segmentation of the frontal lobes is a labour intensive process. It requires that many manual judgements based on established anatomy are made regarding lobar boundaries. This increases the potential for measurement error and subsequent variability. One potential solution to this problem is to transform regional templates onto the MRI scans of each subject in order to calculate the regional volumes.

Post mortem studies confirm that the frontal cortex is commonly involved in PSP, contributing to the neuropsychological deficits that are seen in this disease (Cordato *et al.*, 2000). The results of MRI studies suggest that the reduced frontal lobe volume can be detected during life and that there are associations with the typical behavioural changes characteristic of PSP (Cordato *et al.*, 2002; Cordato *et al.*, 2005). These studies suggest that frontal lobe volumes measured using MRI during life can differentiate PSP from PD and healthy controls. VBM studies suggest that while frontal lobe volume is reduced in MSA and PSP (Brenneis *et al.*, 2003; Brenneis *et al.*, 2004), there is no clear indication as to whether regional cortical atrophy can differentiate them.

The aim of this section of the thesis was to investigate whether application of a similar quadrant based protocol had utility in measuring frontal and posterior-inferior regional volumes in PSP, MSA-P and PD. A quadrant-based method of regional segmentation has been proposed as a guide to assessment of frontal lobe volume. These methods are far less labour intensive than other manual segmentation techniques and have been applied successfully in Alzheimer's disease (AD) and frontotemporal dementia (FTD) (Chan *et al.*, 2001a).

6.6.2. Methods

A standard anatomic space defined by the MNI 305 average brain (Mazziotta *et al.*, 2001) was split into four sub-volumes. An initial left to right split was made by dividing the whole brain volume along the interhemispheric fissure. These two sub-volumes were each then divided mid-way along the anterior posterior axis to define left frontal and right frontal quadrants. The left and right posterior quadrants were then divided at the midpoint of the superior to inferior axis to define the left and right posterior-inferior volumes. These regions were then transformed onto the baseline scan of each subject (registered to standard space) and thresholding applied.

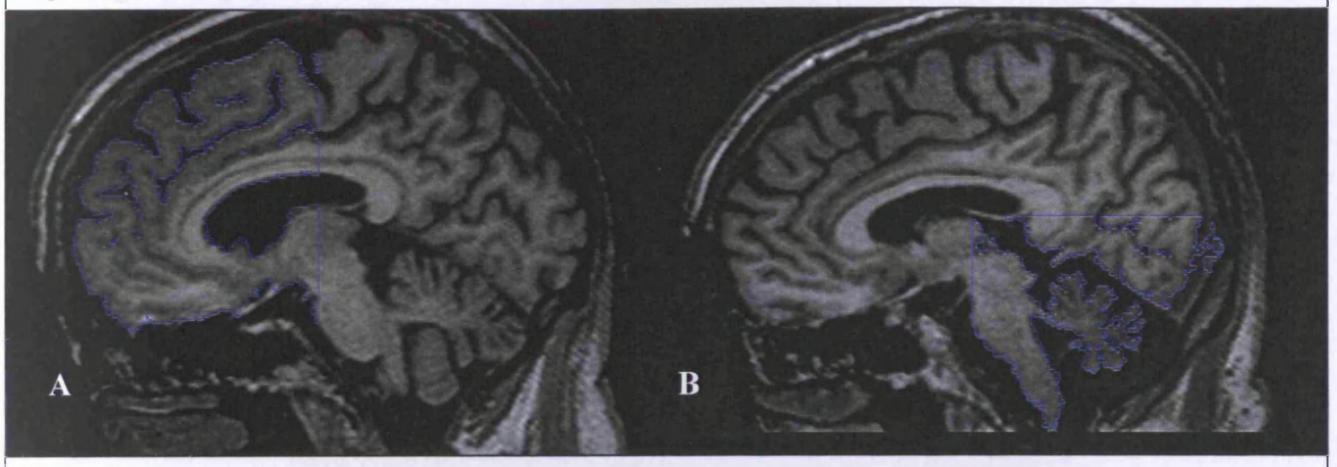
For all scans the MIDAS software tool (Freeborough *et al.*, 1997) was used to delineate the brain/CSF boundary applying a consistent lower threshold of 70% of mean brain intensity to exclude voxels with a lower intensity that were predominantly cerebrospinal fluid (CSF).

In this way, left and right frontal quadrants and left and right posterior inferior regional volumes were determined. Examples of these can be seen in figure 6.15. 3D reconstructions of these volumes are shown in figure 6.16.

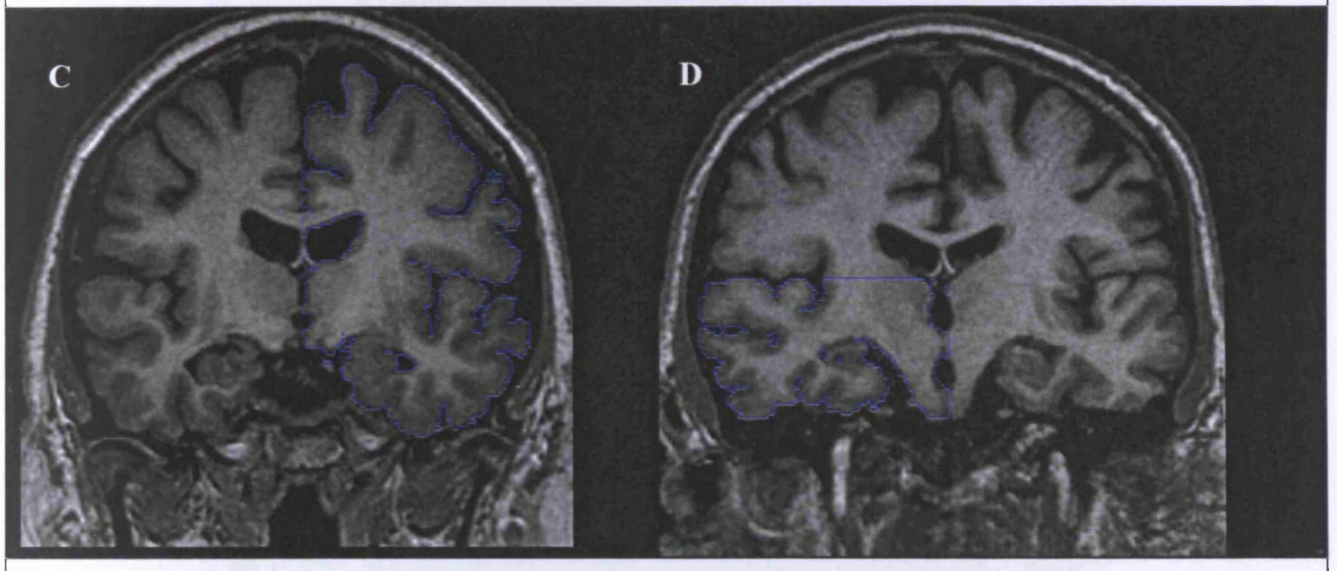
Left and right anterior quadrant volumes and left and right PI volumes were added to give a total frontal volume and total PI volume. This technique does not rely on any manual editing of the ROI.

TIV corrections were applied using the same methods described previously.

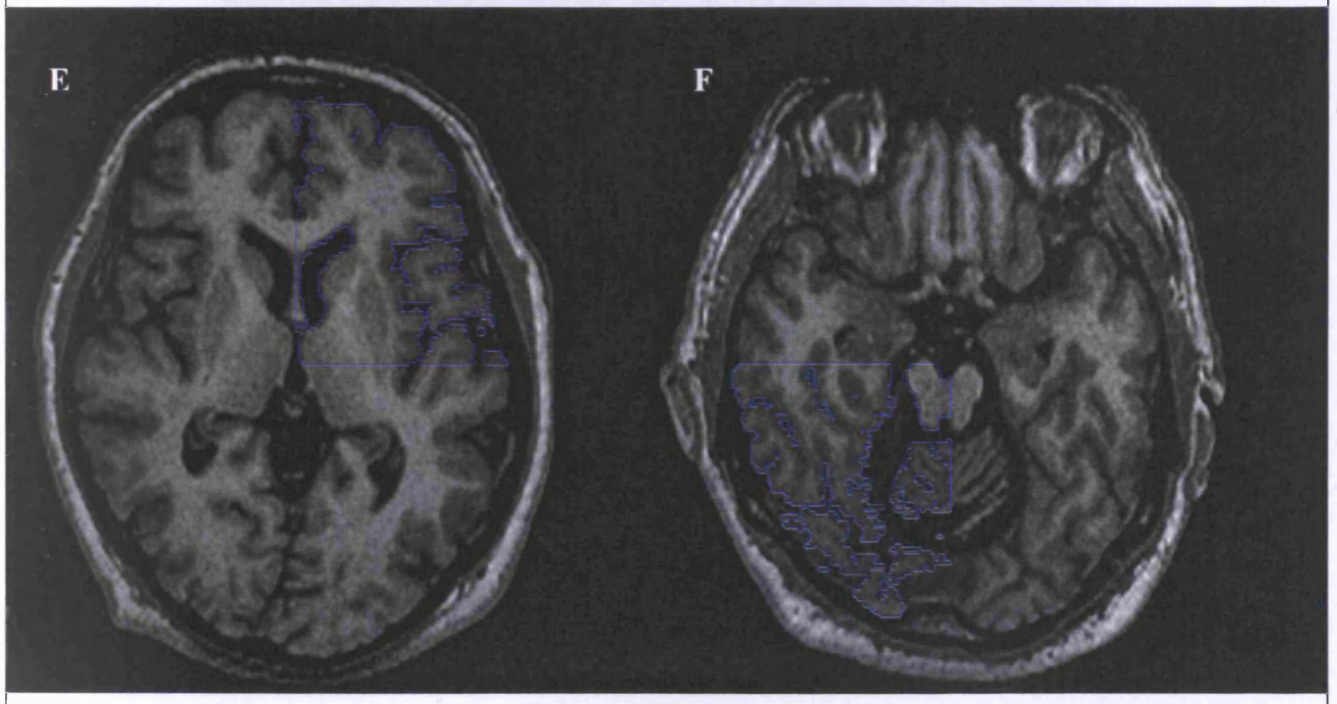
Figure 6.15. *Regional template segmentation of anterior quadrant and posterior-inferior regions.*



Sagittal view showing (A) Left anterior quadrant PSP; and (B) Right postero-inferior region MSA-P

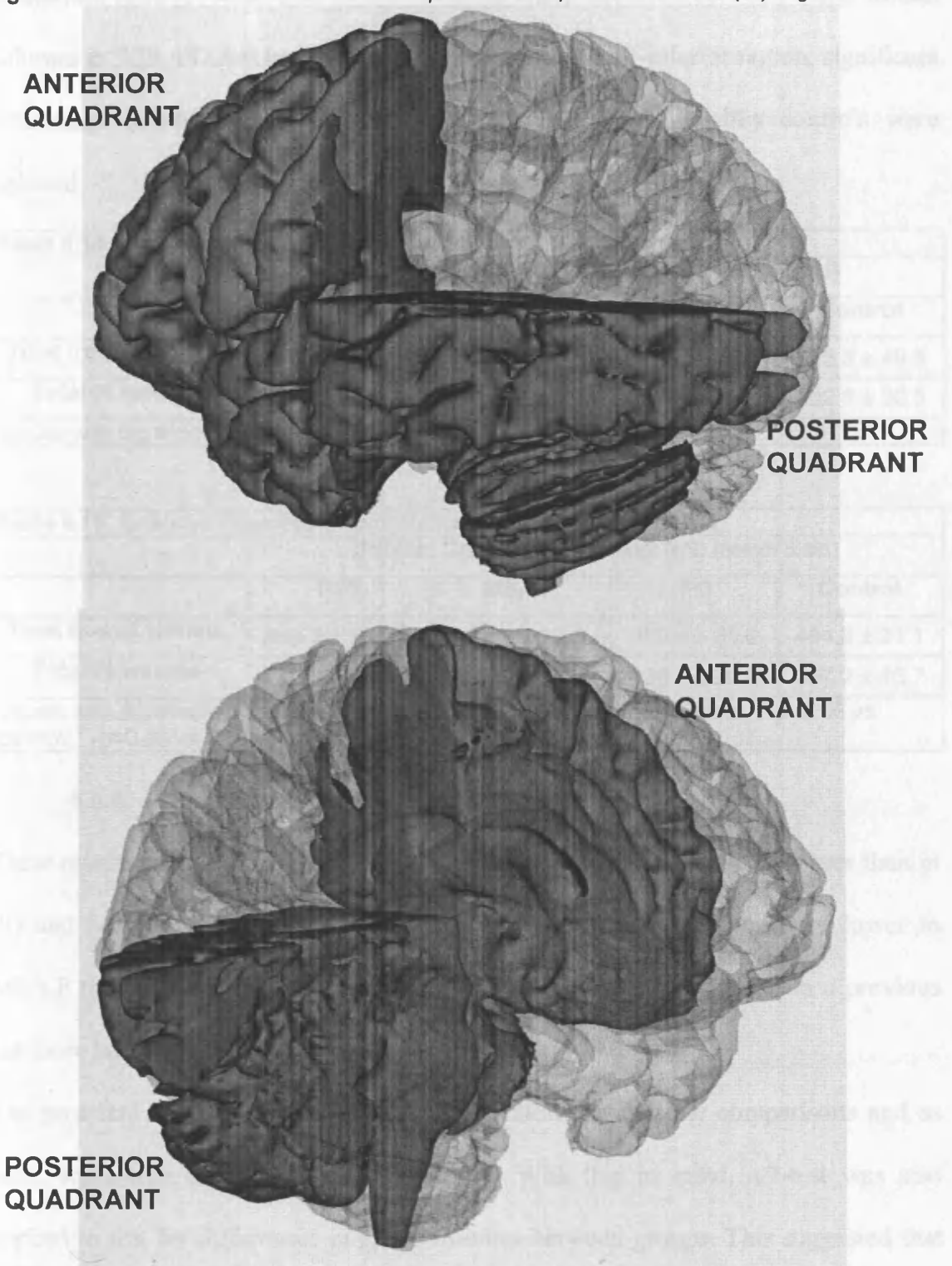


Coronal view showing (C) Left anterior quadrant PSP; and (D) Right postero-inferior region MSA-P



Axial view showing (E) Left anterior quadrant PSP; and (F) Right posterior-inferior region MSA-P

Figure 6.16 *3D reconstructions of anterior quadrant and posterior inferior (PI) region.*



Lateral (above) and medial (below) views of anterior quadrant and posterior inferior (PI) regions.

6.6.3. Results

The frontal quadrant and posterior-inferior regional volumes are shown in table 6.14 (uncorrected) and 6.15 (TIV corrected). One-way analysis of variance with post hoc

Bonferroni tests revealed that significant differences occurred between mean frontal volumes in PSP, PD and healthy controls. For the posterior-inferior region, significant differences between group mean volumes in MSA-P and healthy controls were detected.

Table 6.14. *Quadrant basedROI segmentation volumes (uncorrected)*

	Subject Group ROI volumes (ml) (mean ± sd)			
	PSP	**MSA-P**	**PD**	**Control**
Total frontal volume	457.6 ± 65.1	458.4 ± 44.3	482.1 ± 65.7	475.3 ± 49.8
Total PI volume	326.0 ± 37.7	**311.1 ± 23.0**$^{\blacktriangleleft}$	333.7 ± 39.7	332.3 ± 30.3

Values with significant differences are shown in bold. $^{\blacktriangleleft}$-p=0.04 vs controls.

Table 6.15. *Quadrant basedROI segmentation volumes (TIV corrected)*

	Subject Group ROI volumes (ml) (mean ± sd)			
	PSP	**MSA-P**	**PD**	**Control**
Total frontal volume	**449.2 ± 26.8**$^{\lozenge\blacktriangleright}$	474.8 ± 22.3	480.4 ± 35.0	484.2 ± 21.1
Total PI volume	323.0 ± 21.6	**314.9 ± 22.6**$^{\blacktriangleleft}$	333.9 ± 20.8	337.2 ± 15.7

Values with significant differences are shown in bold. $^{\lozenge}$-p=0.001 vs control, $^{\blacktriangleleft}$-p=0.05 vs control, $^{\blacktriangleright}$-p=0.03 vs PD.

6.6.4. Conclusions

These results suggest that frontal lobe volumes in PSP are significantly lower than in PD and healthy controls and that posterior inferior regional volumes are lower in MSA-P than in healthy controls. These findings support those detailed in a previous but more labour intensive study (Cordato *et al.*, 2002).

The post hoc Bonferroni test is a strict correction for multiple comparisons and as such, some true differences may be missed. With this in mind, a t-test was also applied to test for differences in mean volumes between groups. This suggested that frontal volume in PSP is significantly lower than in MSA-P (p=0.01) as well as PD and healthy controls. The same test was applied to the posterior-inferior measurements, suggesting that this volume was lower in PSP than healthy controls (p=0.01), and lower in MSA-P than in PD (p=0.04).

This is a new and significant finding, as until now, volumetric studies have not demonstrated any difference in frontal volumes between PSP and MSA-P. These easy to apply measures of lobar volume can detect group differences between PSP, MSA-P, PD and healthy controls.

The frontal quadrant includes parts of the anterior temporal lobes as well as the frontal lobes and in addition the method of segmentation means that parts of the posterior frontal lobe are not included and these are affected by the pathological process at post mortem in PSP. The temporal lobes may also be damaged in PSP and may contribute to the reduced frontal volume. For this reason any conclusions regarding clinical associations with frontal lobe volume loss need to be made cautiously.

The posterior inferior volume is made up of the cerebellum and the brainstem as well as the inferior part of the occipital lobe. It is therefore not surprising that the MSA-P group have the lowest mean volume, as the mean cerebellar volume was lower in MSA-P than the other groups. This raises the possibility that this rapid method of regional measurement could provide an approximation of the cerebellar volume. This is supported by the results of a linear regression of cerebellar volume on posterior-inferior volume showing a significant association (R^2=0.55, $p<0.001$).

While the group means may be significantly different, there is considerable variation in the individual volumes measured in all these structures and because of this there is group overlap, hence the clinical utility of these regional volumes may be limited. This is discussed further in Chapter 6.7.

Finally, it is possible that in making assumptions about which regions to study, regions where significant group differences do occur are being missed. Voxel based morphometry offers a method of looking for group differences without making a priori assumptions regarding where they might occur.

6.7. Clinical utility of cross sectional regional assessments

Entering all of the ROI showing significant group differences into a forward stepwise regression analysis revealed the following: midbrain volume alone differentiated PSP from all the other groups ($p<0.0001$, sensitivity 72.2%, specificity 91.9%). The discriminating ability of the MRI findings was improved by adding SCP volume into the analysis ($p<0.0001$, sensitivity 83.3%, specificity 91.9%). The best discrimination of PSP from the other groups was achieved by adding frontal, third ventricle and whole brain volumes, as additional separate variables ($p<0.0001$, sensitivity 88.9%, specificity 97.3%).

Discriminating PSP and just MSA-P on the basis of midbrain volume alone was unreliable ($p=0.004$, sensitivity 83% but specificity 33%) with six false positives (66%). Considering SCP volume, midbrain volume, pons and cerebellar volume together however correctly classified 93% of cases ($p<0.0001$, sensitivity 94.4%, specificity 88.9%). Reduced cerebellar and pontine volume alone were very specific for MSA-P (100%) but not sensitive (44%).

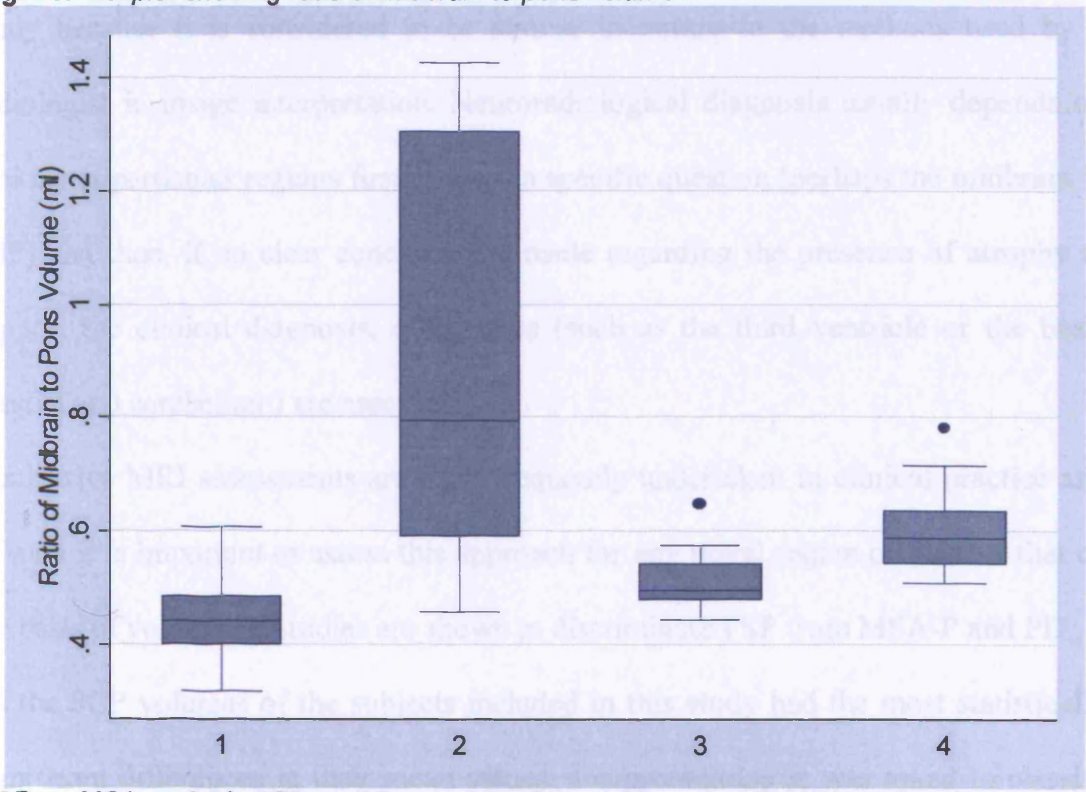
These results suggest that midbrain, superior cerebellar peduncle, pons and cerebellar atrophy considered together rather than in isolation are the best discriminators of PSP from MSA-P. Taken together with the high false positive rate in the MSA group when considering midbrain size alone as a marker of PSP, this suggests that a visual assessment of midbrain to pons ratio might help to better identify cases of PSP. A paper confirming this assumption on the basis of area measurements has recently been published (Oba *et al.,* 2005). The authors demonstrated no overlap between the ratios of midbrain to pons area measurements in PSP and MSA or PD and MSA suggesting that the ratio of midbrain to pontine area measurements might have genuine clinical

utility. Another advantage of this kind of measurement is that it requires no correction for individual variation in baseline brain volume (such as correcting for TIV).

The kind of accumulated tissue loss demonstrated in most cross sectional studies is unlikely to be present early in the course of the disease. This means that there is overlap in regional volume measurements between disease and healthy control groups. So while apparent volume reduction in a region may help to support a particular diagnosis, sensitivities and specificities of 100% will never be achieved.

Based on the volume measurements in this study, similar midbrain to pontine ratios were calculated but as can be seen in the boxplot below (figure 6.17), there is still some group overlap, even though statistically significant differences were detected between the groups using a Mann-Whitney U test.

Fig.6.17 *Boxplot showing ratio of midbrain to pons volume*



PSP vs. MSA, p=0.03; PSP vs. PD, p=0.01; MSA vs. PD, p=0.002; MSA vs. healthy controls, p=0.005.
1-PSP, 2-MSA, 3-PD, 4-healthy controls

The overlap detected in this study could be as a result of using volume rather than area measurements. The area measurements undertaken by Oba *et al* were made on the mid-sagittal T1 weighted image. This image is particularly appropriate for viewing the reduction in midbrain height and it may be that using a volume measurement incorporates tissue that is not so badly affected and hence the ratio is not so robust in terms of discriminating the groups.

6.8. Differential diagnosis using subjective assessment of the SCP

Quantitatively measuring and comparing these regions is one way of studying their ability to differentiate the diseases concerned. Quantitative measurements are important and have definitive clinical utility if they can be shown to discriminate a disease with 100% sensitivity and specificity.

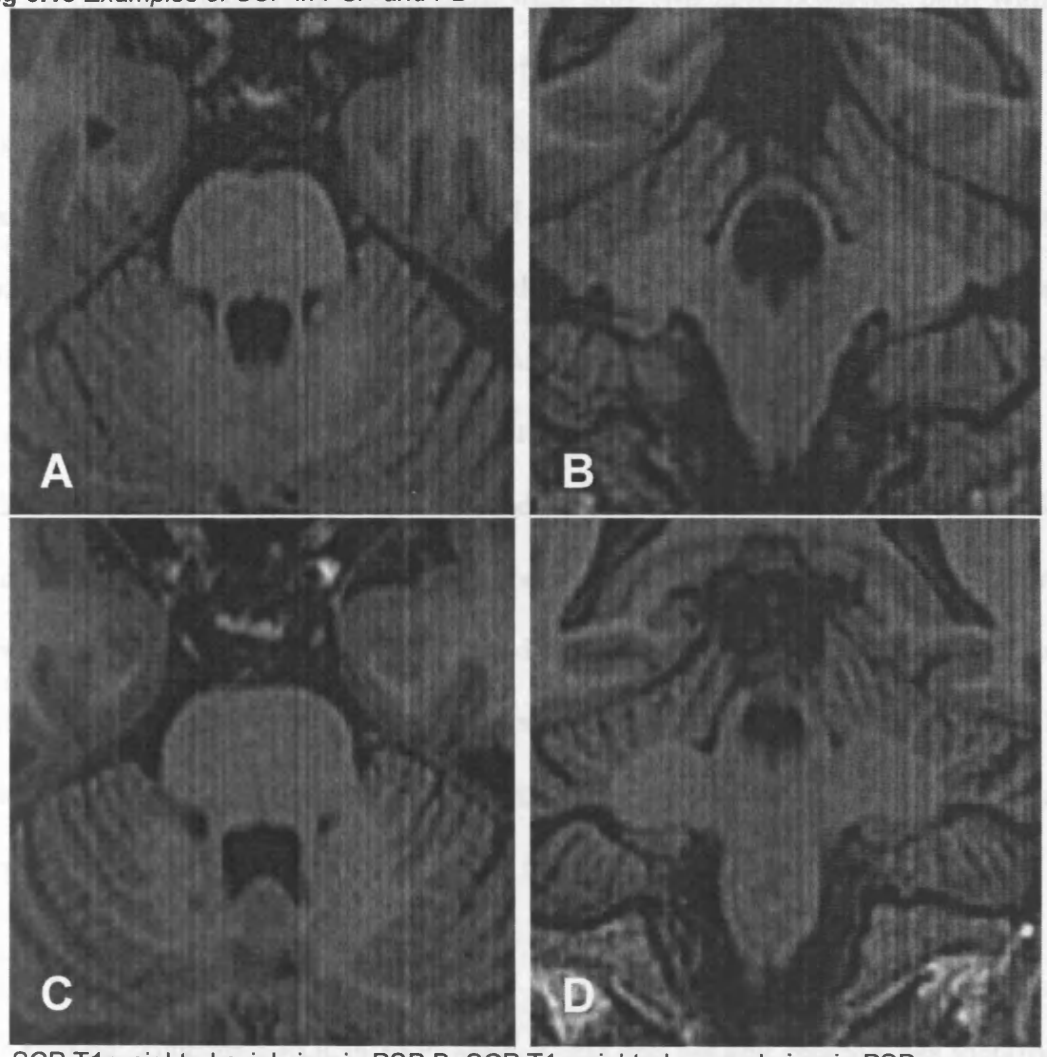
The stepwise logistic regression model was applied as a method of analysis in this study because it is considered to be similar in nature to the methods used by a radiologist in image interpretation. Neuroradiological diagnosis usually depends on looking at particular regions first if asked a specific question (perhaps the midbrain in PSP) and then, if no clear conclusion is made regarding the presence of atrophy to support the clinical diagnosis, other areas (such as the third ventricle or the basal ganglia and cerebellum) are assessed.

Qualitative MRI assessments are most frequently undertaken in clinical practice and as such it is important to assess this approach for any novel region of interest that on the basis of volumetric studies are shown to discriminate PSP from MSA-P and PD.

As the SCP volumes of the subjects included in this study had the most statistically significant differences in their mean values, a neuroradiologist was asked to classify subjects as PSP or non-PSP (MSA-P, PD or healthy controls) on the basis of the presence or absence of atrophy. The neuroradiologist was blinded to the clinical

diagnosis and the results of volume measurements. 19 patients with PSP, 10 with MSA-P, 12 with PD and 12 healthy controls were available for review. More subjects with PSP, MSA and PD were available than for volumetric segmentation as those without a follow up scan were included in this qualitative assessment. Examples of images are shown in figure 6.18.

Fig 6.18 *Examples of SCP in PSP and PD*



A- SCP T1 weighted axial view in PSP B- SCP T1 weighted coronal view in PSP
C- SCP T1 weighted axial view in PD D- SCP T1 weighted coronal view in PD

SCP atrophy on visual rating alone differentiated PSP from other neurodegenerative conditions and controls with a sensitivity of 74% and specificity of 94%. Fourteen of 19 PSP cases were graded as having SCP atrophy while the SCP was rated as atrophic

in only 2 of 34 non-PSP cases. Both non-PSP cases identified as having SCP atrophy had a clinical diagnosis of MSA-P (positive and negative predictive values 88% and 86%). Hence, an easily clinically applicable visual assessment differentiated PSP from the other bradykinetic rigid syndromes with positive and negative predictive values approaching 90%. This is in fact better than the values achieved for this structure alone using logistic regression.

Of the 53 scans visually assessed, there were only two false positives, both with a clinical diagnosis of MSA-P. Three of the five false negative cases had a diagnosis of clinically probable rather than clinically definite PSP, suggesting that SCP atrophy is most likely to be seen in PSP with typical clinical features. The false negative cases did not have the largest SCP volumes of the PSP group on formal measurement nor did the two false positive cases (both MSA-P) have the smallest SCP volumes. These results suggest that on the basis of cross sectional imaging, a number of features in addition to midbrain atrophy can be used to support a clinical diagnosis of PSP. Additionally, visual assessment by an experienced rater of some of these structures appears to be clinically useful. This method was in fact able to discriminate PSP from other diseases better than a volume measurement of the structure. This may be because other regional information is taken into account at the time of a visual assessment, and it is difficult to control for this possibility.

6.9. Overall conclusions

Comparing volumetric ROI in PSP, MSA-P, PD and healthy controls has identified that PSP and MSA-P have differences in regional brain volume which distinguished them from PD and healthy controls and more importantly from each other. The clinico-radiological correlations for these regions are discussed in chapter 12.

There is however considerable overlap in these ROI measurements and this, together with the time taken, the software required, and the corrections necessary for inhomogeneity and TIV, limits their usefulness in clinical practice. Combining measurements in a ratio (such as midbrain to pons) may be more helpful.

A priori assumptions regarding which regions to study are based principally on established knowledge from previous imaging studies, theoretical assumptions about which areas are affected pathological studies. Quantitative and qualitative analysis of the SCP in this study demonstrate that it has clinical utility in discriminating PSP from MSA-P, PD and healthy controls.

Further imaging analysis studies using techniques that do not require assumptions about which regions to study would help to confirm the results outlined in this Chapter and may identify other regions of interest. Voxel based morphometry is one such technique.

7. Cross Sectional Voxel Based Morphometry

7.1. Introduction

Global and regional brain volumes in PSP, MSA-P and PD have so far been considered. Some predictions as to the likely topography of the regional volume loss that can be detected during life on MRI scans can be made on the basis of knowledge from previous imaging studies and from pathological studies. These a priori assumptions have limitations however.

It is reasonable to assume that the earliest pathological changes in PSP might occur in regions functionally responsible for balance and intact gaze. However, volumetric studies carried out so far have required a priori decisions as to which substructures should be assessed.

The size of unambiguous structures with sharp contrast boundaries such as the lateral ventricles are relatively easy to assess. Other morphometric features of disease may be more difficult to quantify by inspection, meaning that many structural lesions of a degenerative disease may be overlooked.

An alternative approach is to use automated techniques that can accurately localize regional volume differences in an unbiased manner. One such approach involves assessment of volume difference between groups of individuals, which can be compared using statistical parametric mapping (SPM), a well validated automated technique for performing such group comparisons.

SPM can be used to identify structural disease related change in brain tissue when applied to volumetric scans. The voxel based morphometry (VBM) approach is not biased and provides an assessment of anatomical differences, between subject groups, throughout the brain (Ashburner and Friston, 2000). VBM removes positional and

global volume differences through spatial normalisation and then detects differences in tissue density by comparing local intensities of tissue after smoothing. VBM avoids some of the shortcomings of other methods by being operator-independent and semi-automated.

Image processing can be divided into:

1) Processes that deal with differences in brain shape

2) Processes that deal with differences in the local composition of brain tissue after macroscopic differences in shape have been discounted. This compares images on a voxel basis after deformation fields have been used to spatially normalise the scans.

Hence residual anatomical differences in the normalised data can be partitioned.

VBM involves spatially normalising all the images to the same stereotactic space, extracting grey matter from the normalised images, smoothing and then performing a statistical analysis to localise and make inferences about group differences. The output shows where tissue concentration differs significantly between groups. Several steps are involved.

1. Spatial Normalisation: Transforms all subjects' imaging data to the same stereotactic space, registering each image to the same template image by minimising the sum of the squared differences in voxel image intensity between them. This method of spatial normalisation does not attempt to match every cortical feature exactly but corrects for global brain shape differences. If the spatial normalisation was exact, then all segmented images would appear identical and no differences would be identified. A mean image of the series (or another registered image) is used to estimate the warping parameters that map it onto a template already conforming to

some anatomical space. Put simply, this finds the deformation that is most likely given the data.

Spatially normalised images should have a relatively high resolution so that tissue-extraction methods are not excessively confounded by partial volume effects. These occur when voxels contain a mixture of different tissue types (e.g. white matter voxels close to CSF voxels).

2. Image partitioning: Spatially normalised images are next partitioned into grey matter, white matter and CSF. The tissue classification method also includes a correction for image intensity non-uniformity.

3. Pre-processing of tissue segments: Grey matter images are smoothed, meaning that each voxel in the smoothed image contains the average concentration of grey matter from around that voxel. The original segmented images contain values between 0 and 1. Most of these values are close to either one or other extreme. Smoothing (e.g. with an 8mm gaussian kernel) results in the data becoming more normally distributed, increasing the validity of the parametric test. Where possible, the size of the smoothing kernel should be comparable to the size of the expected regional difference between the groups of brains.

4. Statistical analyses: Statistical tests are applied to compare and identify differences in tissue concentration between groups as well as to identify tissue concentrations that are related to specific covariates such as disease severity or age. The significance of any difference noted is then ascertained using the theory of Gaussian random fields (Friston *et al.*, 1996). A voxel wise statistical parametric map comprises the results of many statistical tests, and it is necessary to correct for these multiple dependent comparisons.

The technique also makes some important assumptions.

a) Segmentation must correctly identify grey matter and white matter.

b) Confounding effects must be eliminated. For example, the same MRI scanner and the same MR sequence should be used.

5. Evaluation of segmentation: Segmentation methods require good contrast between different tissue types. However many central grey matter structures have image intensities that are almost indistinguishable from white matter, so tissue classification is difficult in these regions.

The model assumes that all voxels contain only one tissue type. Voxels that contain a mixture of tissue types (e.g. white matter and CSF) may not be modelled correctly. In particular, voxels at the interface between white matter and CSF in the lateral ventricles often appear as grey matter, suggesting tissue differences where there are none.

7.2. Methods

20 patients with PSP, 11 with MSA-P, 12 with PD and 12 healthy controls were included in the cross sectional VBM study.

7.2.1. VBM analysis

The digitised structural MR images were processed on Sun workstations (Sun Microsystems Inc, Mountainview, CA) using Matlab (Mathworks, Natick, Massachusetts, USA) and SPM99 (Wellcome Department of Cognitive Neurology, ION, London). VBM was carried out using an optimised method (Good *et al.*, 2001). First, scans were spatially normalized to a customized template to remove any bias potentially caused by gross structural differences between the groups, which may have resulted in spurious areas of difference being identified. All scans were spatially normalized to the SPM T1 template, then smoothed with an 8-mm, full-width half-

maximum (FWHM) smoothing kernel, followed by averaging to create a customized template. The images were then segmented into grey matter (GM) and white matter (WM), using customized prior probability maps. These segments were then modulated to correct for volume changes occurring during spatial normalization. The images were then smoothed, using an isotropic 8-mm FWHM Gaussian kernel.

Regionally specific differences in grey and white matter were assessed statistically using a one-tailed test to highlight decreases in grey (or white) matter in PSP subjects, compared with MSA-P, PD and controls. Age and sex were included as nuisance variables.

Significance levels were set at $p<0.001$ uncorrected over the whole brain volume. Comparisons of voxels over the whole brain involve a great many tests, and correction for multiple comparisons is desirable. This imposes stricter criteria on each voxel, reducing the number of false-positives. The analysis was repeated at $p<0.05$ corrected for multiple comparisons.

Inverse one-tailed tests were also performed for each segment to ensure that changes seen were due to real disease effects and not movement of artefact.

The analysis was run including GM/WM/ globals, as appropriate, as a covariate, to only highlight differences over and above the general tissue change.

7.3. Results

The clinical characteristics of the patient and healthy control groups are detailed in table 7.1.

Table 7.1 *Mean (SD) Characteristics of patient groups in cross sectional VBM analysis.*					
	PSP	**MSA-P**	**PD**	**HC**	**p**
n	20	11	12	12	
Age	66 (6.1)	62 (7.7)	65.5 (9.2)	67.4 (4.7)	0.3
Disease duration	4.5 (1.5)	5.4 (1.6)	13.3 (6.7)	-	PD vs. PSP/MSA-P, p<0.001
MMSE	26.3 (3.0)	26.4 (3.2)	27.7 (2.5)	29.5 (0.7)	HC vs. PSP, p<0.001; vs. MSA-P, p=0.002; vs. PD, p=0.03
FAB	12.2 (3.0)	14.7 (2.5)	16.5 (1.4)	-	<0.001
UPDRSII	19.6 (6.4)	25 (7.0)	13.9 (5.1)	-	<0.001
UPDRSIII	20.5 (7.6)	26.8 (9.7)	16.7 (5.1)	-	0.01
HY	3.5 (0.7)	3.9 (0.8)	2.8 (0.6)	-	0.001

All results are for analysis carried out using GM/WM globals as covariates and age and sex as nuisance covariates.

Comparing PSP with healthy controls (figure 7.1), revealed reduced grey matter in the region of the midbrain and thalamus, the insular cortex, the caudate nucleus and the medial frontal gyrus. White matter differences were identified predominantly in the region of the cerebral peduncles. After correcting for multiple comparisons, grey matter tissue differences remained in the insular cortex and white matter tissue differences remained in the region of the upper midbrain and cerebral peduncles.

Comparing PSP with PD (figure 7.2), revealed grey matter tissue differences in the region of the superior cerebellar peduncle and white matter tissue differences confined to the midbrain and cerebral peduncles. Only the white matter tissue differences in the region of the upper midbrain survived correction for multiple comparisons.

Figure 7.3 shows tissue differences between PSP and MSA-P demonstrating reduced grey matter volume in the region of the dentate nucleus of the cerebellum. There were no differences in white matter and the grey matter tissue differences did not survive correction for multiple comparisons.

Comparing MSA-P and healthy controls demonstrated reduced grey matter in the putamen bilaterally and in the cerebellar grey matter as well as in the region of the

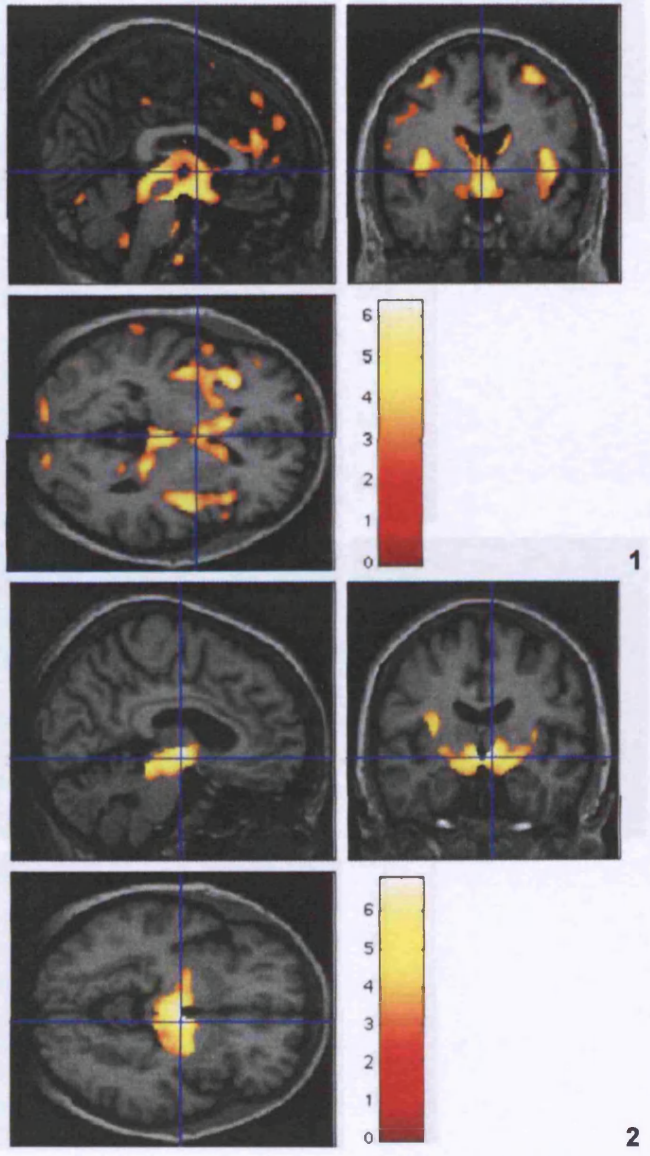
posterior thalamus. Reduced grey matter in the insular cortex bilaterally was also detected. White matter tissue differences were detected in the pons and the middle cerebellar peduncle and these remained after correction for multiple comparisons (figure 7.4).

Figure 7.5 shows grey matter tissue differences in MSA-P compared to PD, present in the putamen and the cerebellar grey matter. White matter tissue differences were present in the pons and middle cerebellar peduncle (figure 7.6) but only the differences in the pons survived correction for multiple comparisons.

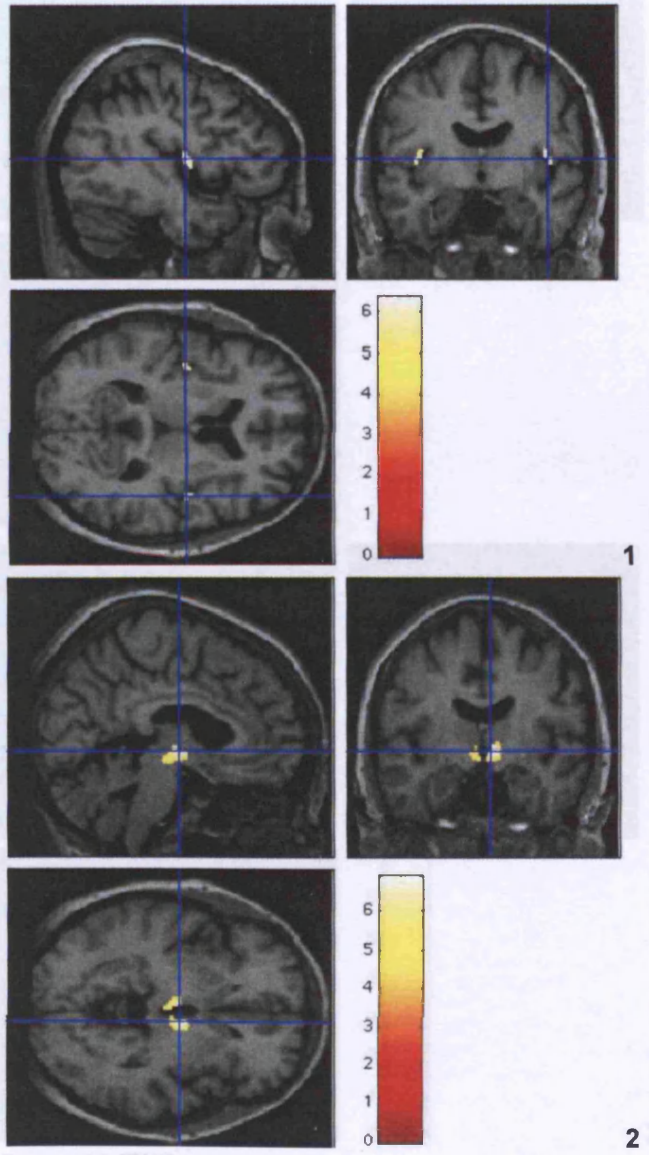
Comparing MSA-P and PSP (figure 7.7) revealed tissue differences to be present only in the grey matter of the putamen and the white matter of the pons. None of these regions survived correction for multiple comparisons.

Finally, when comparing PD and healthy controls grey matter differences were detected in the region of the hypothalamus and occipital gyri. White matter differences were minimal and neither survived correction for multiple comparisons (figure 7.8).

Figure 7.1 Patterns of reduced grey and white matter in PSP compared with healthy controls.
Statistical parametric maps (SPM) are shown indicating regions of reduced grey (1) and white (2) matter voxel intensity in PSP compared with healthy control subjects. In PSP, significantly ($p<0.001$) reduced grey matter intensity is present in the midbrain, cerebral peduncles and the insular cortex as well as the frontal cortex. White matter intensity is reduced in the cerebral peduncles. A strict correction for multiple comparisons (B) reveals significant differences between PSP and healthy controls in the insular cortex and midbrain.

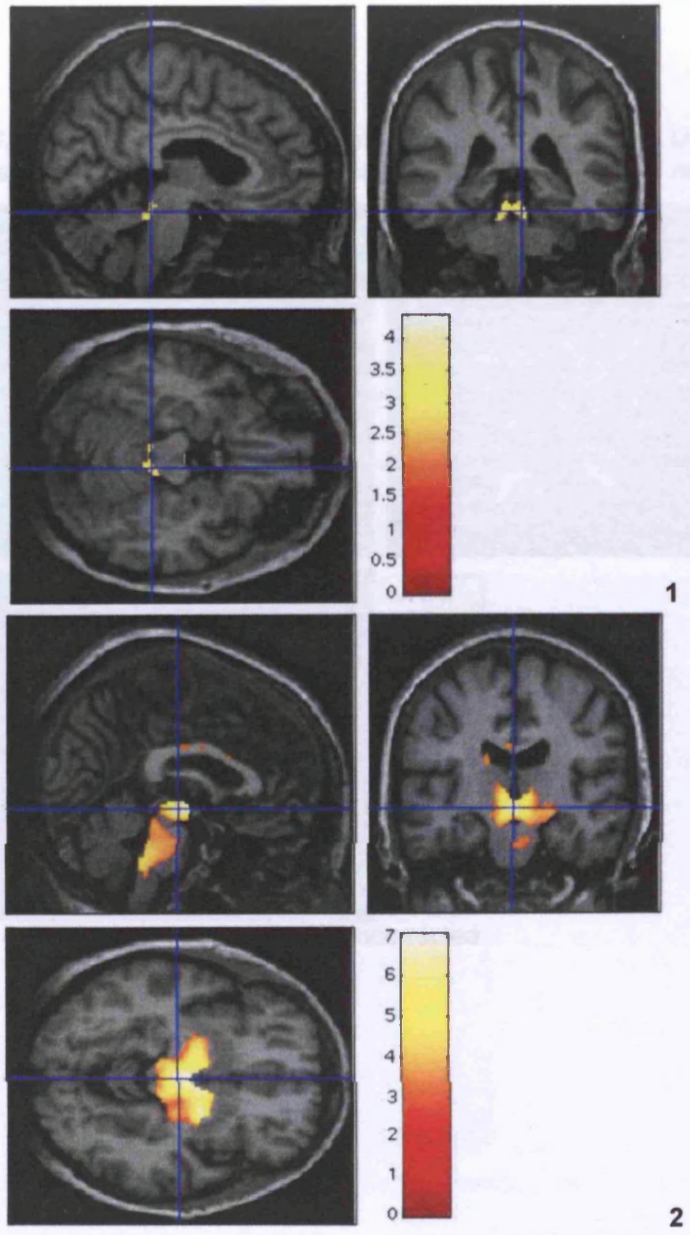


A $p<0.001$ uncorrected, 1-grey matter; 2-white matter

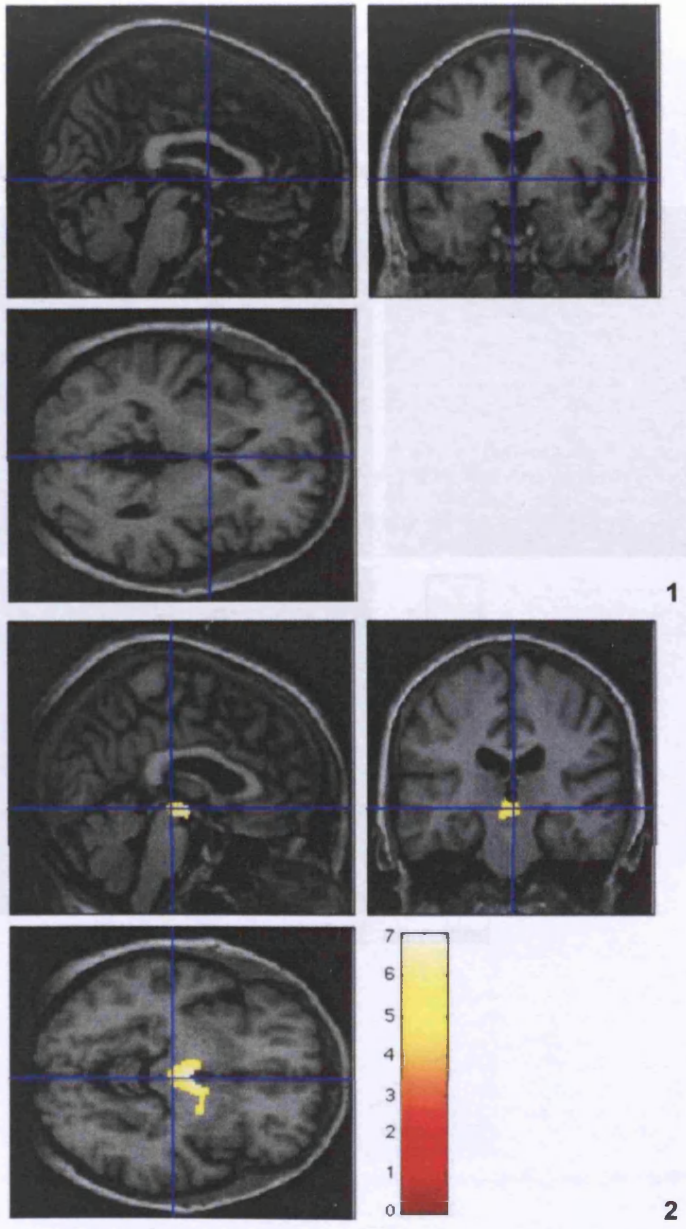


B $p<0.05$ FWE corrected, 1-grey matter; 2-white matter

Figure 7.2 Patterns of reduced grey and white matter in PSP compared with PD.
Statistical parametric maps (SPM) are shown indicating regions of reduced grey and white matter voxel intensity in PSP compared with PD.

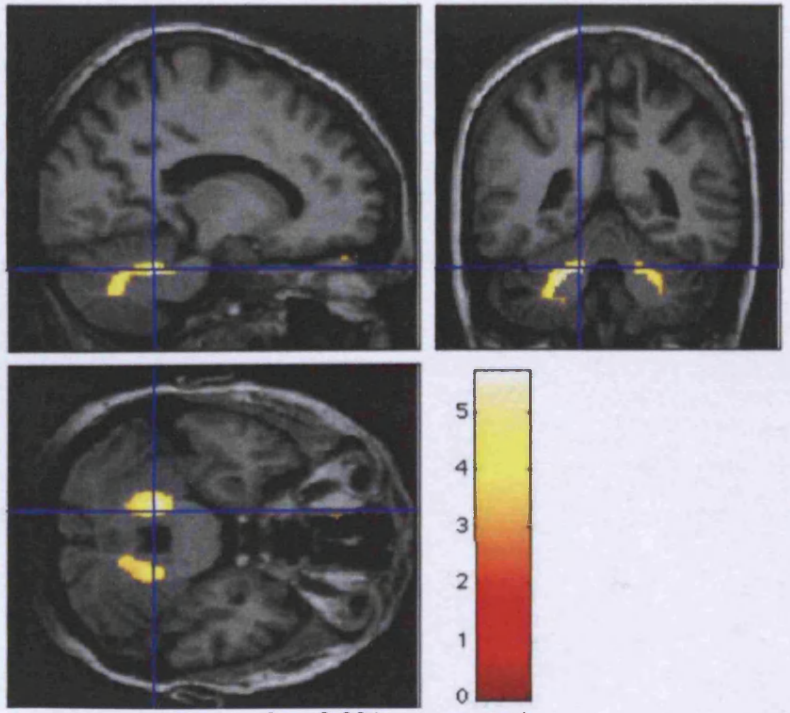


A p<0.001 uncorrected, 1-grey matter; 2-white matter

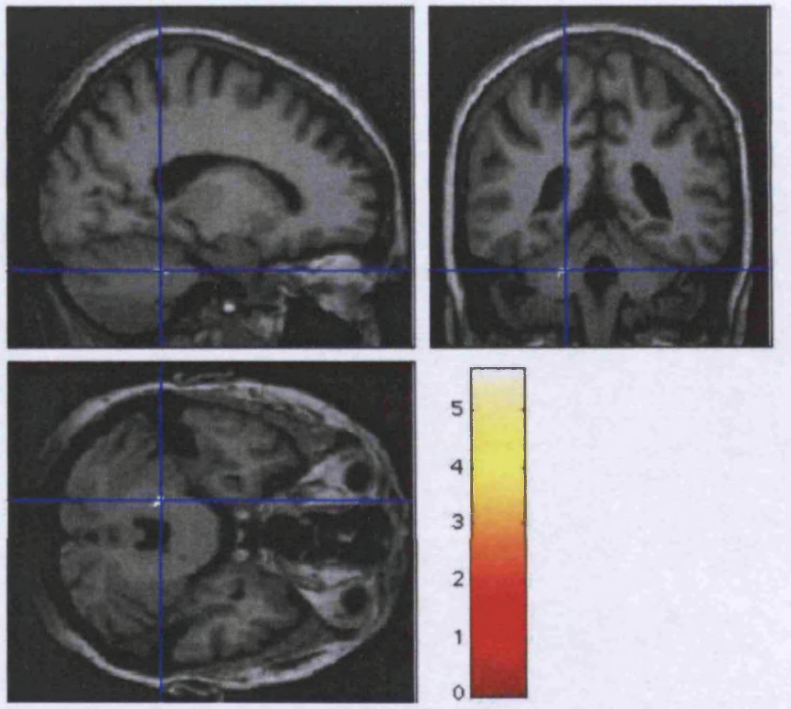


B p<0.05 FWE corrected, 1-grey matter; 2-white matter

Figure 7.3 Patterns of reduced grey matter in PSP compared with MSA-P.
Statistical parametric maps (SPM) are shown indicating regions of reduced grey matter voxel intensity in PSP compared with MSA-P.

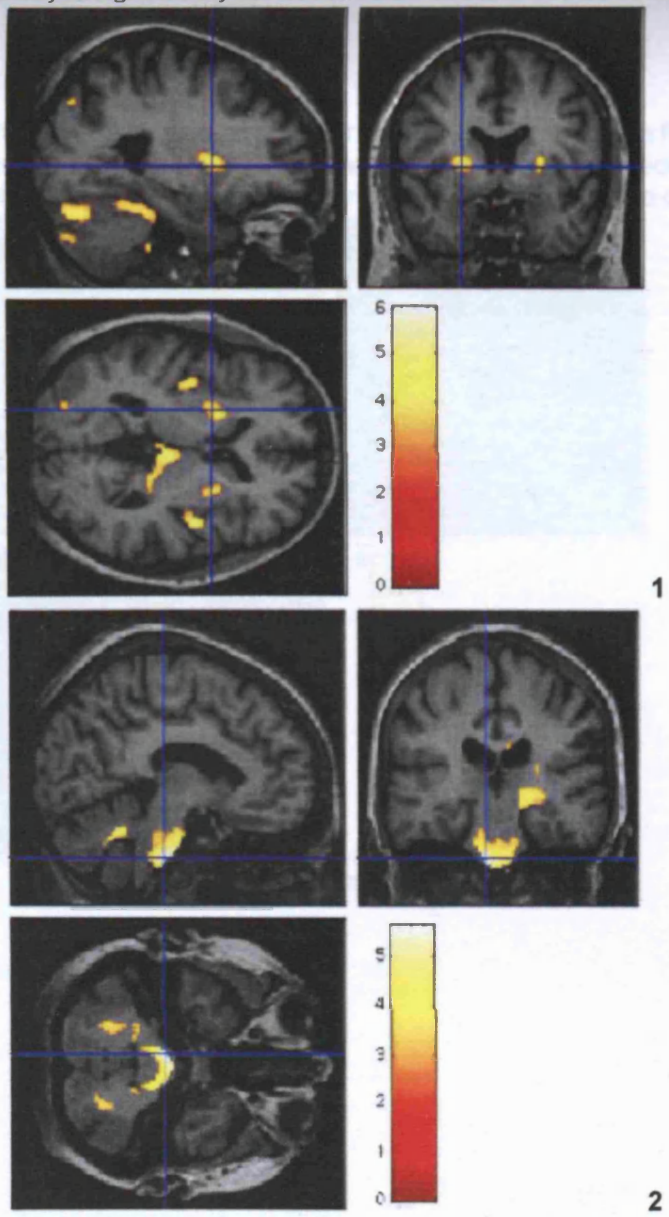


A p<0.001 uncorrected

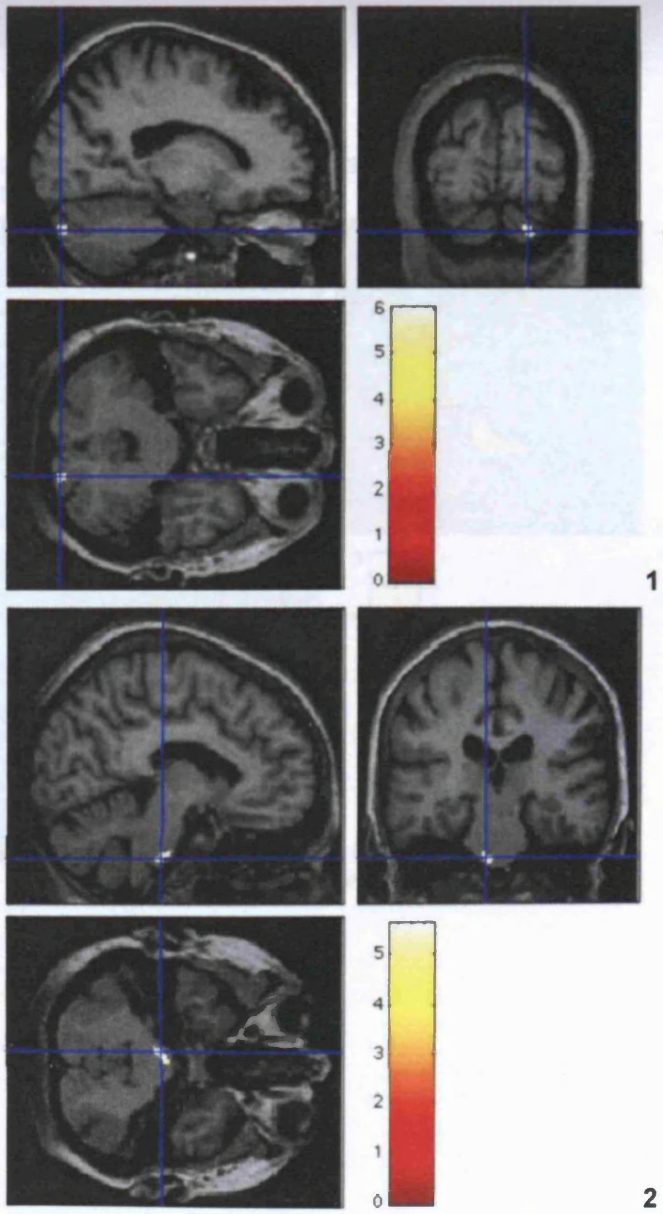


B p<0.05 FWE corrected

Figure 7.4 *Patterns of reduced grey and white matter in MSA-P compared with healthy controls.* Statistical parametric maps (SPM) are shown indicating regions of reduced grey and white matter voxel intensity in MSA-P compared with healthy controls. In MSA-P significantly reduced grey matter voxel intensity is detected in the cerebellar cortex and in the region of the putamen. White matter tissue intensity is significantly different in the Pons and middle cerebellar peduncles.

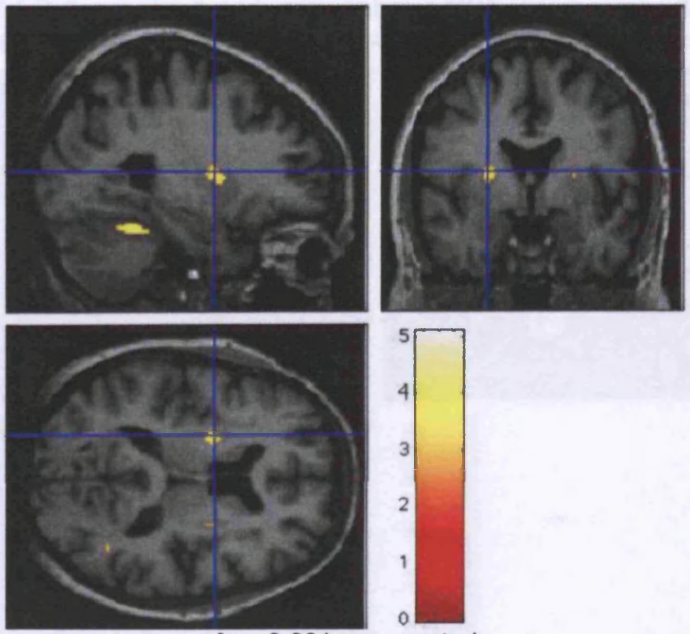


A p<0.001 uncorrected, 1-grey matter; 2-white matter

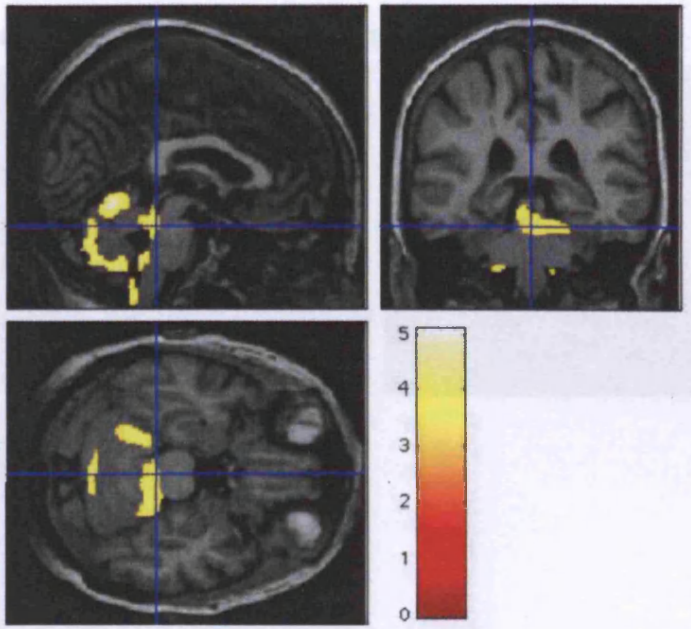


B p<0.05 FWE corrected, 1-grey matter; 2-white matter

Figure 7.5 Patterns of reduced grey matter in MSA-P compared with PD.
Statistical parametric maps (SPM) are shown indicating regions of reduced grey matter voxel intensity in MSA-P compared with PD. Regions of reduced voxel intensity are demonstrated in the putamen (A) and the cerebellar grey matter (B).

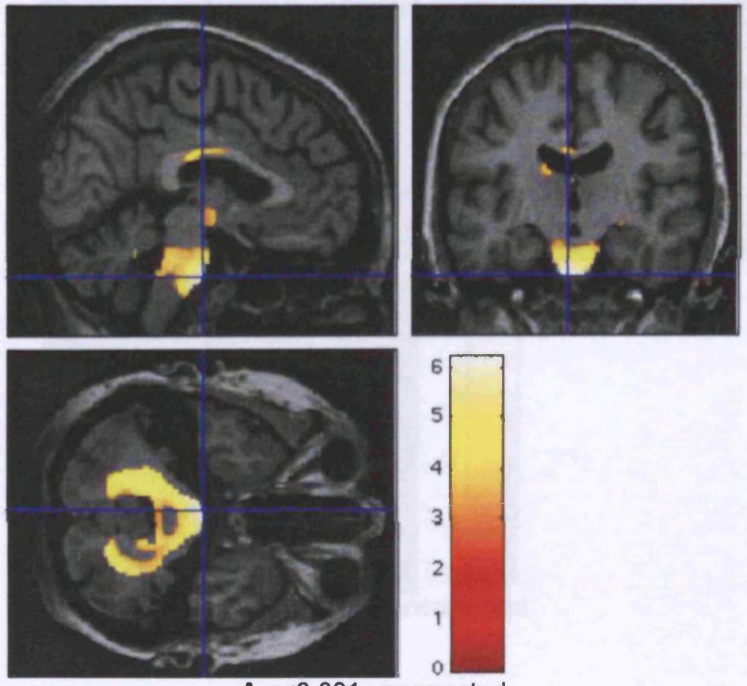


A p<0.001 uncorrected

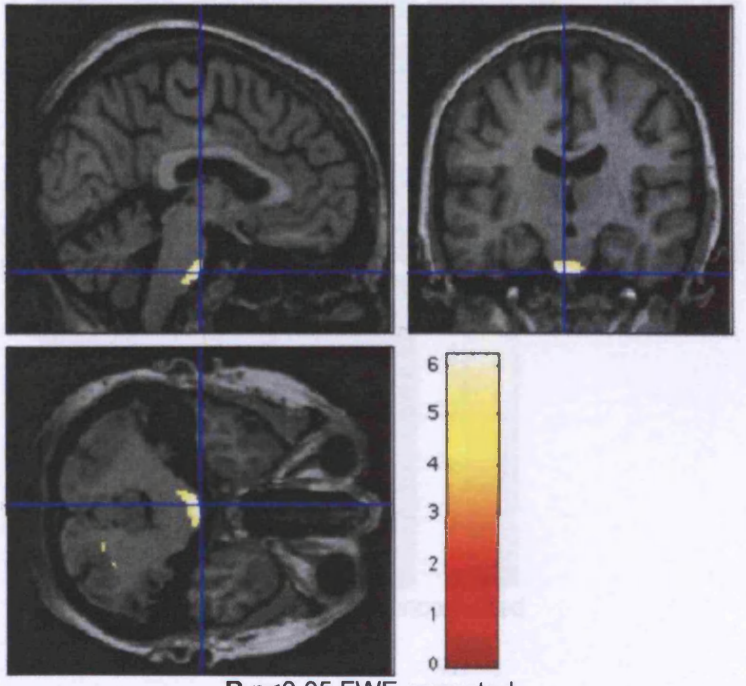


B p<0.001 uncorrected

Figure 7.6 Patterns of reduced white matter in MSA-P compared with PD.
Statistical parametric maps (SPM) are shown indicating regions of reduced white matter voxel intensity in MSA-P compared with PD.

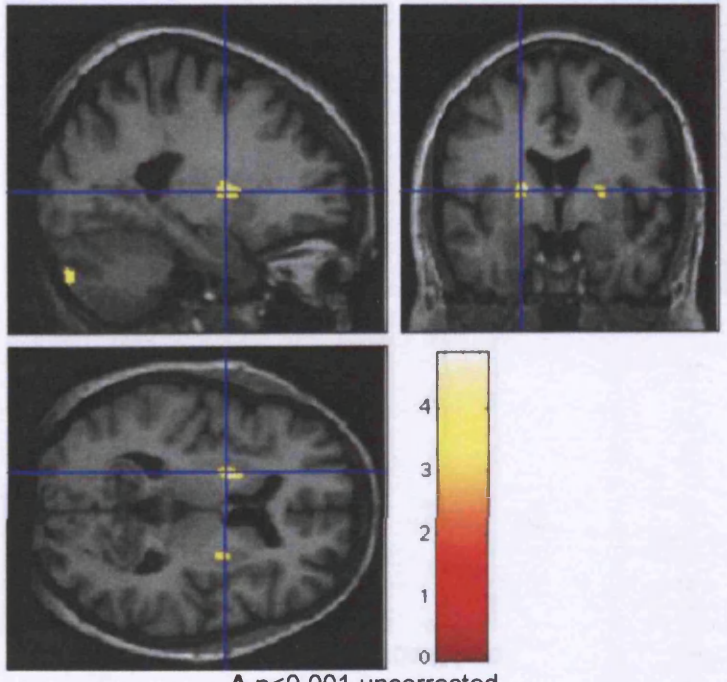


A p<0.001 uncorrected

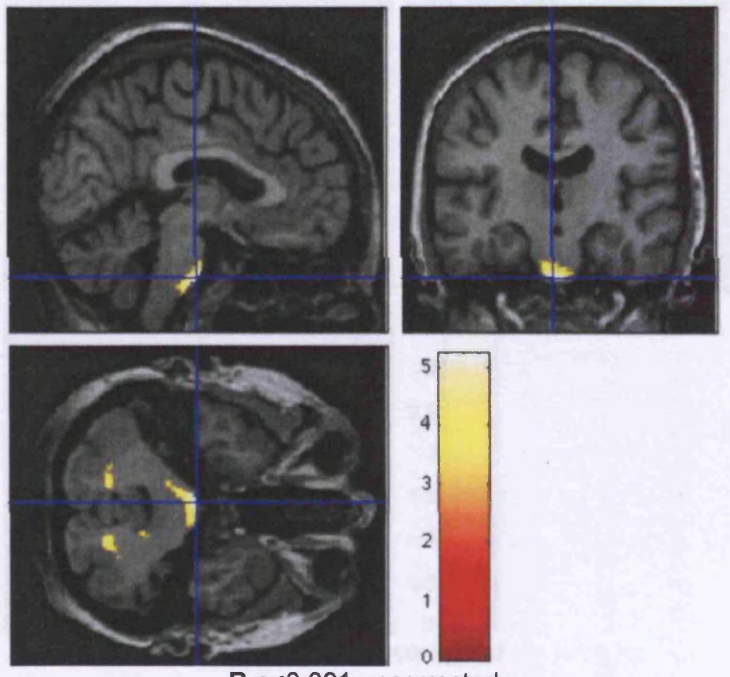


B p<0.05 FWE corrected

Figure 7.7 Patterns of reduced grey and white matter in MSA-P compared with PSP.
Statistical parametric maps (SPM) are shown indicating regions of reduced (A) grey and (B) white matter voxel intensity in MSA-P compared with PSP.

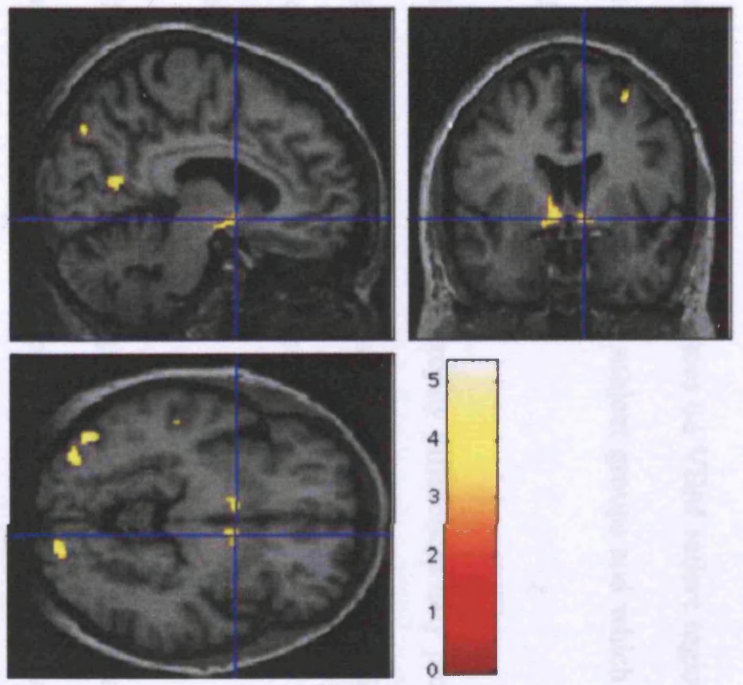


A p<0.001 uncorrected

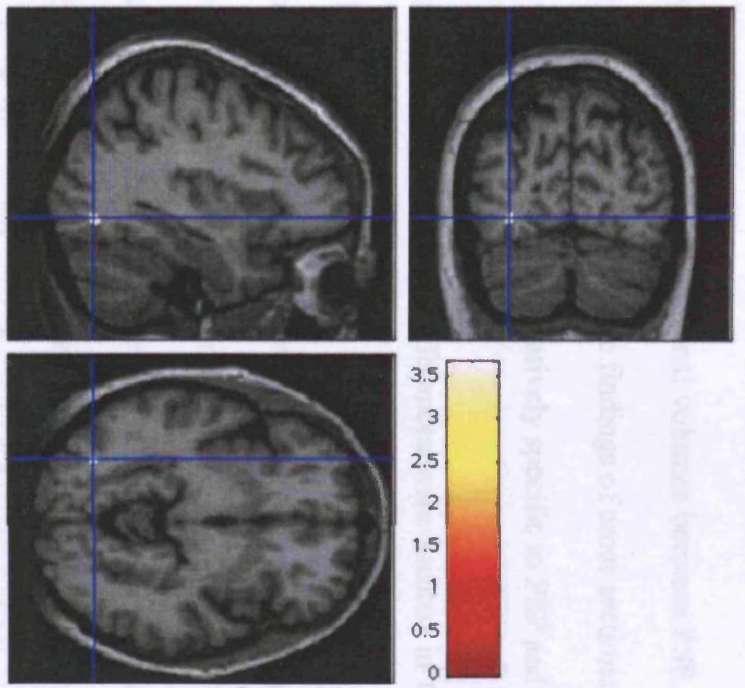


B p<0.001 uncorrected

Figure 7.8 Patterns of reduced grey and white matter in PD compared with healthy controls.
Statistical parametric maps (SPM) are shown indicating regions of reduced (A) grey and (B) white matter voxel intensity in PD compared with healthy controls.



A p<0.001 uncorrected



B p<0.001 uncorrected

7.4. Conclusions

VBM revealed differences in regional volumes between PSP, MSA-P and PD as well as healthy controls and supports the findings of cross sectional ROI analysis that these regional volume differences are relatively specific to PSP and MSA-P.

In PSP these significant group differences were located in the upper brainstem and cerebral peduncles and in MSA-P, in the putamen, pons and cerebellum. In addition, this method of analysis involving no a priori assumptions regarding regional volume differences, detected significant group differences between PSP and healthy controls with reduction in grey matter in the insular cortex, the caudate nucleus and the medial frontal gyrus, confirming that frontal cortical volume loss is present in PSP. This supports the findings of the quadrant based region of interest analysis described in chapter 6.

The statistically significant differences on VBM reflect regional atrophy patterns that are consistently different between subject groups and which cannot be explained by global differences.

Using VBM to identify regional atrophy differences may prove to be of more use in clinical practice than making assumptions regarding regional atrophy based on postmortem studies as changes seen at post mortem largely reflect the disease in its latest stages.

No significant differences in the grey matter or white matter structure of the upper brainstem was seen between PSP and MSA-P, but did show significant reduction in pontine and putamen volume in MSA-P. It is surprising that no difference in midbrain

volume between PSP and MSA-P was detected but the region of interest segmentation described in the previous chapter identified midbrain volume loss in MSA-P as well as in PSP. It was reassuring that this unbiased method of analysis produced very similar results.

In PSP compared with MSA-P, grey matter volume differences were identified in the region of the dentate nucleus of the cerebellum. This deep cerebellar nucleus is severely damaged in PSP, but not in MSA-P. Efferent fibres from this nucleus travel in the, the dentato-rubro-thalamic tract, which forms the superior cerebellar peduncle (Tsuboi *et al.,* 2003). Atrophy of this region in PSP compared to healthy controls and PD is also demonstrated using the VBM methods described in this chapter, supporting the findings of the ROI studies already outlined.

In MSA-P, significant grey matter differences were detected in the putamen. This region was not assessed in the ROI analysis described in Chapter 6. It is well established that putaminal damage is usually severe in MSA-P, and that a more severe lesion in this region as opposed to the cerebellum results in the more parkinsonian phenotype of MSA-P (Ozawa *et al.,* 2004). Qualitative evaluation of MRI can be applied to help discriminate MSA-P from PD on the basis of putaminal hypointensity/hyperintensity (Schrag *et al.,* 2000). VBM analysis confirmed regional differences between MSA-P and PSP, PD and healthy controls in this deep grey matter structure. Cerebellar grey matter differences were also identified in the MSA-P subjects confirming that as suggested by the results detailed in Chapter 6, the cerebellum is atrophied in MSA even when parkinsonism predominantes.

VBM has already been used to study degenerative basal ganglia disorders including PD, dementia with Lewy bodies (DLB) (Burton *et al.*, 2002), and MSA (Brenneis *et*

al., 2003; Specht *et al.*, 2003). Recently 2 studies in PSP (Brenneis *et al.*, 2004; Cordato *et al.*, 2005) have also been published. In PD patients without dementia, reduced grey matter volume in the frontal lobe has been demonstrated (Burton *et al.*, 2004). In the same study, grey matter atrophy was noted in the occipital lobe when comparing PD with dementia to PD without dementia. The mean MMSE in the PD group in the present study did not suggest marked cognitive impairment but some occipital grey matter tissue differences compared with healthy controls were detected, while no frontal volume differences were found.

In MSA-P, VBM has previously detected differences in cortical and subcortical regions compared to controls (Brenneis *et al.*, 2003), suggesting that a degree of cortical atrophy occurs relatively early in the disease. This may help to explain why no differences in cortical volume were seen between PSP and MSA-P, even though more extensive cortical atrophy probably does not occur until later in the disease process in MSA (Konagaya *et al.*, 1999b; Wakabayashi *et al.*, 1998).

Cerebral and subcortical metabolic differences have been demonstrated in PSP with reduced glucose metabolism in frontal cortical areas as well as midbrain and basal ganglia regions (Hosaka *et al.*, 2002). This corresponds well with the known areas of pathological involvement (Litvan *et al.*, 1996b). Whether this regional hypometabolism predates histopathological change is unclear.

Pathological volumetric studies comparing regional brain atrophy in PSP and Lewy body disease are limited. One study suggested that PSP could be differentiated by marked atrophy of the internal globus pallidus and that a trend for greater frontal atrophy correlated with dementia in PSP (Cordato *et al.*, 2002).

Given the abnormalities seen in the frontal cortex in metabolic imaging and more recently in other VBM studies, it is perhaps surprising that the results in this chapter do not show a more dramatic difference between the patient groups in frontal cortical regions, particularly between PSP, PD and MSA-P was not found here. There is a significant difference between the FAB scores in the two groups, and frontal cortical atrophy is common in PSP at autopsy. However, frontal cortical tissue differences between PSP and healthy controls were clear and of a similar distribution to those demonstrated in a previous study where frontal atrophy was the most striking finding, even though the mean disease duration is slightly longer here than in another recently published study of VBM in PSP in which frontal lobe atrophy was the most dramatic finding (Brenneis *et al.,* 2004). Another MRI study has suggested frontal volume losses occur early in PSP (Cordato *et al.,* 2002) and one of the most recent VBM studies detected an association between frontal cortical volume and the severity of the neuropsychological deficits associated with frontal cognitive impairment (Cordato *et al.,* 2005). Another possibility is that some of the early frontal behaviour seen in PSP occurs as a result of extensive basal ganglia and brainstem pathology.

Since VBM is generating maps of statistical differences in brain tissue, it is important that each of the regions (grey, white, and CSF) is examined in combination to highlight areas where change is due to tissue atrophy rather than movement of structures caused by tissue loss in other areas. By assessing the grey, white, and CSF maps together, a clearer picture can be obtained of where genuine tissue losses have occurred.

As with most cross-sectional studies, the present study is somewhat limited by the inherent differences in regional volumes between individuals. Calculation of atrophy

rates, from serial MRI scans would overcome some of these problems. However, using reasonably sized subject groups should overcome some of these interindividual differences. This study demonstrates that MRI and VBM can be used to detect regions of interest for further imaging studies and volume analysis in PSP. One such region not evaluated in the volumetric studies detailed in chapter 6 is the middle cerebellar peduncle in MSA-P. This would be a candidate region to study with novel imaging techniques such as diffusion weighted MRI (see chapter 8).

In conjunction with postmortem studies, MRI information may help us to determine the sequence of cortical and subcortical structures affected in the early and later stages of PSP, potentially aiding earlier diagnosis and identification of MRI markers of disease progression. However cross sectional studies can only imply atrophy and no firm conclusions as to the ongoing rate of atrophy can be drawn.

8. Cross Sectional Diffusion weighted imaging

8.1. Introduction

Diffusion refers to the general transport of matter whereby molecules or ions mix through normal thermal agitation in a random way. Each molecule in a sample behaves independently from the others. Collision between molecules provokes a random displacement of each one without a preferred direction, tracing the path known as the random walk.

Given a time interval, it is possible to calculate a statistical measure of diffusion distance averaged over an equilibrium ensemble of molecules (the root mean square distance), but it is not possible to say how far an individual given molecule has moved during that time.

Ficks first law of diffusion states that: the diffusion process drives tracer from areas of high to areas of low concentration.

$$F = - D (\delta C/\delta X)$$

Where F = the rate of transfer of a diffusing substance through unit area, C = concentration of diffusing substance and X = the space co-ordinate measured normal to the section. D is the diffusion coefficient with units $(length)^2(time)^{-1}$.

Every fluid has a characteristic intrinsic self diffusion coefficient reflecting the mobility of molecules in their environment (Crank, 1998). Magnetic resonance imaging can be made sensitive to dynamic displacements of water molecules over 10^{-8} and 10^{-4}m in a timescale of a few milliseconds to seconds. These magnitudes are similar to the size of cellular structures.

In the adult brain, white and grey matter, have structural complexity, which affects the diffusion coefficient of water in tissues. In bundles of nerve fibres, diffusion of water is easier along the fibres than across them. This dependence on direction is

"anisotropy". Damage to the tissue structure, reduces this anisotropy. Even at an early stage of disease, measurement of the diffusion coefficient may be useful.

Quantification of diffusion behaviour is relatively straightforward, the complexity of brain tissue structure means that the results obtained are dependent on the pulse sequence and acquisition parameters. This means careful planning of imaging is essential.

The diffusion tensor is a measure of diffusion behaviour in all directions. From the tensor, we obtain the mean diffusivity (equivalent to the apparent diffusion coefficient, ADC, over all directions). This increases with tissue destruction.

In MRI, the tissue is excited using a radio frequency (RF) pulse, then the magnetization is refocussed with for example a 180 degree RF pulse at time t = TE/2. This refocusing pulse reverses the phase of the proton spins leading to a cancellation of the phase due to magnetic field inhomogeneity and the formation of a spin echo at t =TE. If there has been a drift of the spins between dephasing and rephasing, then the refocusing of the spins is only partial and there is an extra damping (loss) term in the description of transverse magnetisation evolution. The resulting phase distribution is not coherent, resulting in decreased signal amplitude at t = TE.

The amount of diffusion attenuation depends on two factors;

1. The sample structural characteristics, which determine the motion of the spins.
2. Sequence parameters, which determine the magnetic field gradients and the time during which the diffusive motion takes place.

On DWI, areas of higher diffusivity (such as the ventricles) appear darker. Areas with similar diffusion properties in every direction are said to be isotropic and on DWI

have the same signal characteristics irrespective of the direction of application of the diffusion gradient.

Areas where tissue structure favours water movement along a particular direction are characterised by different diffusion coefficients in different directions. In these cases, the signal attenuation reflects the diffusion properties in the direction of application of the diffusion gradients. These areas are anisotropic.

When diffusion gradients are applied along the gradients of highest diffusivity, signal attenuation is most evident.

DWI is open to qualitative evaluation (for example in acute cerebrovascular disease). It is sensitive to the choice of acquisition parameters and positioning in the scanner. The MRI signal also depends upon the T2 properties of the tissue and signal increase due to T2 may be interpreted as a reduction in diffusivity, or may mask an increase in diffusivity (the T2 shine through effect).

An ADC map shows the average diffusion coefficient of the tissue contained in that voxel, measured along the direction of application of the diffusion gradient.

Diffusion weighted MR imaging has the potential to detect ultrastructural changes earlier than T1 or T2 weighted MRI. Increased ADC measurements in the presence of normal T2-weighted images have been shown in a number of diseases including MS and AD (Symms *et al.*, 2004). Studies of degenerative bradykinetic-rigid syndromes have established that DWI can help to distinguish MSA from PD and PSP from PD (Ohshita *et al.*, 2000; Schocke *et al.*, 2002; Seppi *et al.*, 2003). So far however, the ability of DWI to discriminate PSP from MSA-P has not been demonstrated. It is important to establish whether ADC measurements correlate with clinical measures of disease severity as this imaging technique may have potential as an outcome measure

in clinical trials. It is possible that DWI may provide information about tissue damage which is complementary to the information obtained from volumetric MRI.

The aim of the DWI analysis was to determine: whether ADCs in regions such as the middle cerebellar peduncle, in which tissue differences between PSP, MSA-P, PD and healthy controls had been identified, could help to further discriminate between PSP and the other conditions; whether the ADC values are associated with clinical markers of the diseases; and whether any conclusions can be drawn from the relationship between the ADC value and ROI volumes described in Chapter 6.

8.2. Methods

MRI Protocol

Axial diffusion weighted MRIs (DWI) were obtained using spin-echo planar sequences (repetition time, TR- 1000ms, echo time, TE- 99ms, matrix size-128 x 128 pixels, field of view, FOV: 240mm x 240mm, slice thickness 7mm and slice gap 2mm). Diffusion-sensitising gradients were applied along 3 orthogonal directions in turn at two gradient strengths corresponding to diffusion-weighting ('b') factors of 0 and 1000 s/mm^2. Following magnitude reconstruction, images obtained with diffusion weighting in the 3 orthogonal directions were co-added yielding diffusion-weighted images, which were essentially rotationally invariant.

T1 images were obtained using a spoiled gradient-echo technique as described before (Chapter 7) (256 x 256 matrix, FOV 24 x 18cm, TI/TR/TE/NEX/FLIP = 650ms/13ms/5.4ms/1ms/35°). This yielded 124 contiguous 1.5mm thick slices.

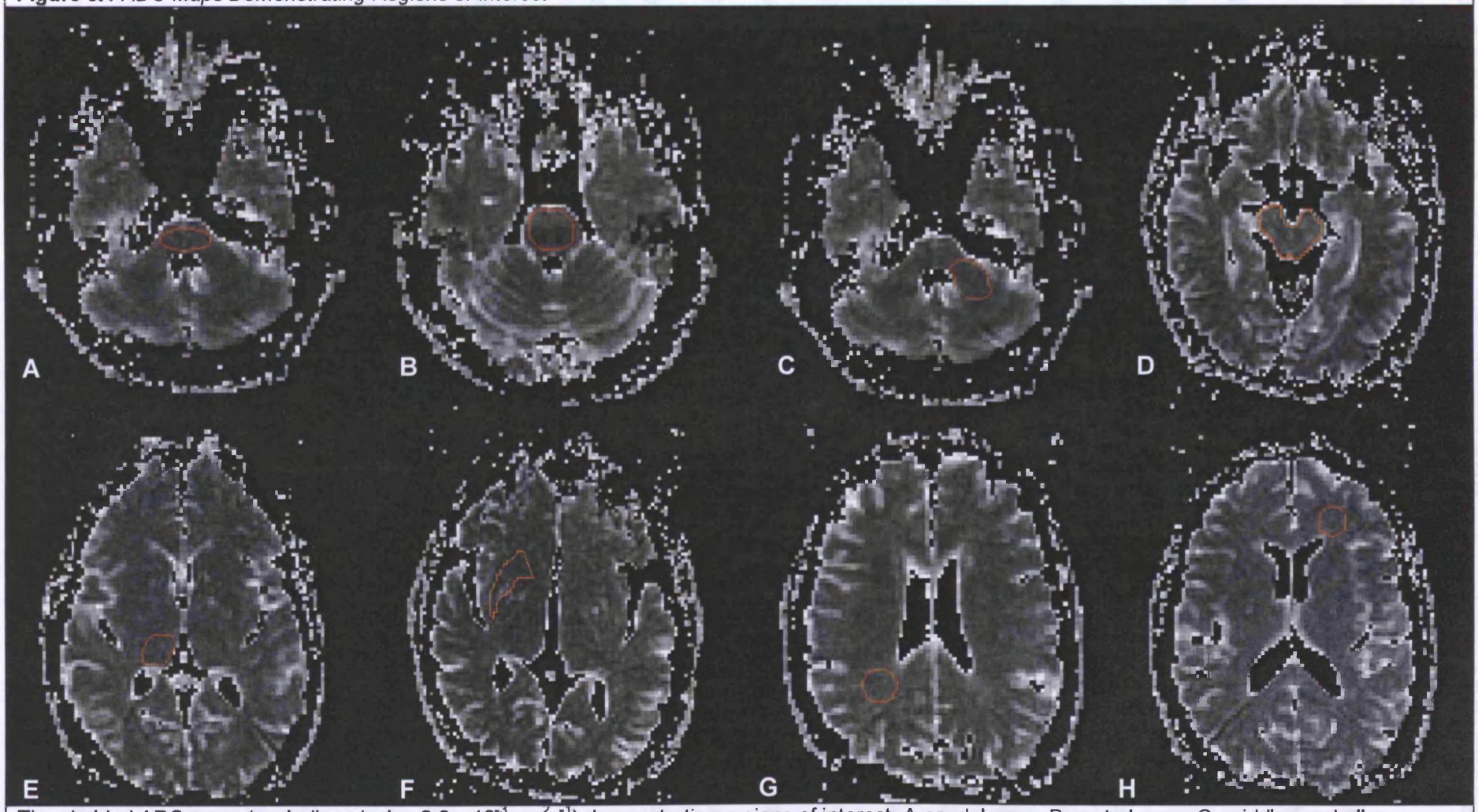
Image analysis

ADC maps were calculated on a pixel by pixel basis assuming a signal dependence of the form $S/S_0 = \exp(-b.ADC)$ where S and S_0 represent the signal intensities with and without diffusion sensitisation. The ADC maps were then thresholded at an arbitrary

level of 2.0 x 10^{-3} mm^2s^{-1} in order to exclude pixels predominantly containing CSF (Seppi *et al.*, 2003). Anatomical regions of interest (ROIs) were manually defined on the (T2-weighted) images obtained with zero diffusion weighting, using DispImage software (Plummer, 1992). A single rater blinded to the identity of each subject analysed all the images. Elliptical ROIs (figure 8.1) were placed in the left and right middle cerebellar peduncles (MCP), the caudal and rostral pons, the decussating fibres of the superior cerebellar peduncle, the left and right thalamus, the splenium and genu of the corpus callosum, the left and right caudate nuclei, putamen and globus pallidus, and the left and right frontal and parietal white matter as well as the left and right centrum semi-ovale. The midbrain was manually outlined at the level of the collicular plate.

The ROIs were then transferred to the inherently co-registered ADC.

Figure 8.1 *ADC Maps Demonstrating Regions of Interest*



Thresholded ADC maps (excluding pixels >2.0 x $10^{-3}mm^2s^{-1}$) demonstrating regions of interest: A-caudal pons; B- rostral pons; C- middle cerebellar peduncle (MCP); D- Midbrain; E-Thalamus; F-Putamen; G-parietal white matter; H- frontal white matter.

For each ROI, the mean (SD) ADC, minimum ADC and maximum ADC value were recorded.

In 10 randomly selected subjects, rADCs were measured a second time and the intra-rater repeatability (within subject SD) and reliability (intraclass correlation coefficient, ICC) calculated using methods described previously.

Statistical analysis

For non-parametric data, the Kruskall Wallis test was used to compare the mean rADC measurements between PSP, MSA-P, PD and healthy controls. An adjustment for multiple comparisons was made when interpreting the p-values for regions that were not independent of each other. So for the left and right MCP, the adjusted significance level was set at p=0.025. For the SCP and the midbrain, similarly the significance level was set at 0.025. Where significant differences were found (after adjustment for multiple comparisons), multiple group comparisons were performed using Mann-Whitney U tests. Normally distributed data was assessed using a one-way analysis of variance with post-hoc t-tests when significant group differences (after adjustment for multiple comparisons) were found.

To assess the ability of rADC measurements to distinguish PSP and MSA-P, forward stepwise logistic regression analysis was performed including the rADC values that were significantly different between PSP and MSA-P, as independent variables.

For non-parametric data, the Spearman rank test was used to study correlations between disease duration and mean rADC values as well as between motor deficit (UPDRS II, III and Hoehn and Yahr score) and mean rADC values in the brainstem structures for PSP and MSA-P. The same test was applied to study correlations between the MMSE and mean rADC in the parietal white matter and the FAB and

mean rADC in the brainstem structures, thalamus and frontal white matter. For normally distributed data, linear regression models studying the same associations were applied. Age and disease duration were included as covariates.

Correlations between total intracranial volume (TIV) corrected cerebellar volume and mean MCP rADC, midbrain volume and mean midbrain rADC, as well as pons volume and mean pons ADC at both levels of the pons, in PSP and MSA-P were also studied, using multiple regression models with age, disease duration and TIV corrected whole brain volume as covariates. Partial correlation coefficients (Corr.) were calculated to assess the strength of the correlation if the covariates were held constant.

8.3. Results

The diffusion-weighted protocol requires that the subject spends only a very short time in the MRI scanner. This means that some individuals who were unable to tolerate a ten-minute volumetric sequence were able to manage the DWI sequence. For this reason, the number of patients included in the cross sectional DWI data is different from the volumetric data.

20 PSP, 11 MSA-P and 12 PD patients had DWI scans. Only 7 healthy control subjects had DWI scans. Significant differences in disease duration, MMSE score, FAB score, UPDRS II and III as well as Hoehn and Yahr score are reported in table 8.1.

ICC values of 0.9 for the MCP, pons, midbrain and thalamus, and values for the left and right frontal white matter of between 0.7 and 0.9 suggested a reliable measurement method in these areas. The within subject standard deviations in these regions were low indicating good repeatability. Measurements in the corpus callosum,

SCP, CN, putamen and globus pallidus as well as other cerebral white matter regions suggested poor repeatability and reliability (ICC, 0.2-0.6).

Table 8.1 *Mean (SD) clinical characteristics of DWI subjects*

	PSP	MSA-P	PD	HC	P values
n	20	11	12	7	
Age	66.1 (6.1)	62.0 (7.7)	65.5 (9.2)	63.1 (8.6)	0.5
Disease duration	4.6 (1.8)	5.4 (1.6)	13.3 (6.7)*	-	<0.001
MMSE	25.6 (3.2)#	26.4 (3.2)##	27.7 (2.5)	29.3 (1.11)	0.006
FAB	11 (3.7)**	14.7 (2.5)▲	16.5 (1.4)	-	<0.001
UPDRS 2	20.0 (7.0)	24.9 (7.0)	13.9 (5.1)►	-	0.001
UPDRS 3	20.8 (7.8)	26.8 (9.7)	16.7 (5.1)▼	-	0.01
HY	3.6 (6.7)	3.9 (0.8)	2.8 (0.6)◄	-	0.001
Post hoc tests: PD vs. PSP/MSA-p<0.001; #PSP vs. PD-p=0.04, vs. HC-p=0.001; ##MSA vs. HC-p=0.02; **PSP vs. MSA-p=0.006, vs.PD-p<0.001, ▲MSA vs. PD-p=0.05; ► HC vs. PSP –p=0.01, vs. MSA-p<0.001; ▼HC vs. MSA-p=0.005;◄ HC vs. PSP-p=0.003, vs. MSA-p=0.001.					
MMSE-mini mental-state examination; FAB-frontal assessment battery; UPDRS 2/3-unified Parkinson's disease rating scale parts 2 and 3; HY-Hoehn and Yahr score.					

Significant differences in mean (SD) rADC values between the different disease groups are reported in table 8.2.

Table 8.2 *Mean (SD) rADC ($x10^{-3}mm^2 s^{-1}$).*

	PSP	MSA-P	PD	HC	P values
n	20	11	12	7	
L MCP	0.729 (0.059)	0.885 (0.16)	0.710 (0.041)	0.701 (0.038)	<0.001*
R MCP	0.718 (0.044)	0.878 (0.15)	0.714 (0.037)	0.705 (0.023)	<0.001**
Caudal Pons	0.789 (0.089)	0.845 (0.133)	0.771 (0.101)	0.763 (0.079)	0.11
Rostral Pons	0.754 (0.114)	0.872 (0.149)	0.708 (0.061)	0.780 (0.074)	0.005#
Midbrain	0.925 (0.087)	0.907 (0.062)	0.867 (0.053)	0.909 (0.066)	0.19
SCP	0.845 (0.092)	0.793 (0.046)	0.791 (0.034)	0.771 (0.033)	0.06
L Thalamus	0.813 (0.074)	0.766 (0.022)	0.775 (0.028)	0.767 (0.023)	0.09
R Thalamus	0.823 (0.074)	0.801 (0.028)	0.786 (0.037)	0.790 (0.031)	0.35
Splenium CC	0.853 (0.136)	0.800 (0.060)	0.908 (0.245)	0.961 (0.133)	0.16
Genu CC	1.037 (0.191)	0.960 (0.134)	0.962 (0.166)	0.923 (0.087)	0.52
L putamen	0.763 (0.053)	0.756 (0.031)	0.767 (0.069)	0.752 (0.025)	1.0
R putamen	0.780 (0.070)	0.769 (0.041)	0.759 (0.042)	0.760 (0.051)	0.9
L Globus Pallidus	0.782 (0.078)	0.757 (0.029)	0.767 (0.051)	0.771 (0.061)	0.4
R Globus Pallidus	0.795 (0.083)	0.773 (0.055)	0.767 (0.045)	0.748 (0.071)	0.2
L Caudate	0.837 (0.098)	0.811 (0.104)	0.823 (0.060)	0.840 (0.089)	0.8
R Caudate	0.866 (0.098)	0.852 (0.095)	0.840 (0.072)	0.870 (0.077)	0.8
L Frontal WM	0.825 (0.072)	0.799 (0.062)	0.777 (0.036)	0.792 (0.010)	0.14
R Frontal WM	0.833 (0.069)	0.799 (0.050)	0.778 (0.051)	0.802 (0.020)	0.05
L Parietal WM	0.810 (0.056)	0.782 (0.039)	0.774 (0.036)	0.783 (0.026)	0.19
R Parietal WM	0.825 (0.057)	0.801 (0.061)	0.791 (0.050)	0.806 (0.024)	0.3
L Centrum Semi Ovale	0.790 (0.082)	0.743 (0.044)	0.733 (0.035)	0.740 (0.057)	0.05
R Centrum Semi Ovale	0.772 (0.072)	0.734 (0.049)	0.723 (0.033)	0.730 (0.045)	0.04
Post hoc tests: * MSA vs. PSP/PD-p<0.001, vs. HC-p=0.001; ** MSA vs. PSP/PD-p<0.001, vs. HC-p=0.002; # MSA vs. PSP-p=0.017, vs. PD –p=0.002.					

Significant group differences in mean (SD) rADC values were detected in the MCP and rostral pons. Post hoc analysis confirmed that group differences in MCP rADC and pontine rADC arose because values in MSA-P were significantly greater than in PSP, PD and healthy controls. After correcting for multiple comparisons, group differences detected for right frontal white matter (p=0.05) and left and right centrum semi-ovale (p=0.05 and p=0.04) fell short of significance. No other significant group differences were detected.

Mean rADC values (mean of left and right rADC) for regions with left and right values were calculated and are reported in table 8.3 below.

Table8.3 *Mean(SD) combined left and right rADC ($x10^{-3}mm^2 s^{-1}$).*

	PSP	MSA-P	PD	HC	P values
n	20	11	12	7	
MCP	0.723 (0.05)	0.881 (0.160)*	0.712 (0.03)	0.703 (0.03)	<0.001
CN	0.851 (0.09)	0.831 (0.06)	0.831 (0.06)	0.795 (0.08)	0.9
Thalamus	0.818 (0.07)	0.783 (0.02)	0.780 (0.03)	0.735 (0.02)	0.2
Putamen	0.772 (0.05)	0.762 (0.03)	0.763 (0.05)	0.703 (0.04)	0.9
Globus Pallidus	0.788 (0.08)	0.765 (0.04)	0.767 (0.04)	0.855 (0.06)	0.3
Frontal	0.829 (0.07)#	0.799 (0.05)	0.778 (0.04)	0.797 (0.01)	0.01
Parietal	0.817 (0.05)	0.791 (0.04)	0.782 (0.04)	0.795 (0.01)	0.06
Centrum Semi Ovale	0.781 (0.07)@	0.738 (0.05)	0.723 (0.03)	0.735 (0.05)	0.02

Post hoc tests: * MSA vs. PSP/PD/HC-p<0.001; # PSP vs. PD-p=0.005, @PSP vs. MSA –p=0.03, vs. PD-p=0.02.

Significant group differences were detected in the MCP (greater in MSA-P than in PSP, PD and controls -p<0.001), the frontal white matter (greater in PSP than PD-p=0.005) and in the centrum semi ovale (greater in PSP than MSA-P and PD-p=0.02). Forward stepwise logistic regression analysis confirmed that rADC in the right MCP was the most significant region in terms of discriminating MSA-P from PSP (p=0.01). The sensitivity and specificity for rADC in the MCP were calculated using a receiver operator characteristic (ROC) curve analysis. The optimal cut-off level (with an area under the ROC curve of 0.89) to discriminate MSA-P from PSP was a MCP rADC ≥ 0.733 $x10^{-3}mm^2$ s^{-1} (sensitivity of 91% and a specificity of 80%). This correctly classified 84% of PSP and MSA subjects.

There were no correlations between any of the mean rADC values and disease duration or age in any of the disease groups. The Spearman test revealed correlations between the rADC in the rostral pons and motor deficit measured using the Hoehn and Yahr scale in MSA-P (rho=0.61, p=0.05), rADC in the rostral pons and UPDRS 3 in PSP (rho=0.44, p=0.05). There were also associations detected between thalamic rADC and UPDRS III in PSP (left thalamus: rho=0.54, p=0.014; right thalamus: rho=0.052, p=0.018) and between the right thalamus rADC and FAB in PSP (rho= -0.67, p=0.001). The association for the left thalamus approached but did not reach significance (p=0.08).

TIV corrected cerebellar volume was significantly associated with rADC values in the left and right MCP in MSA-P (LMCP: Corr.= -0.81, p=0.05; RMCP Corr.= -0.9, p=0.02; combined left and right MCP, Corr.=-0.86, p=0.03), but not in PSP or PD. There was a trend towards an association between TIV corrected midbrain volume and midbrain ADC in PSP (p=0.06). In MSA-P, a lower TIV corrected pontine volume was associated with increased rADC in the caudal pons (Corr.=-0.82, p=0.05). No associations between pontine volume and rADC in the pons were detected in PSP.

8.4. Conclusions

This study demonstrates that rADC values in the middle cerebellar peduncles and pons are significantly increased in MSA-P compared to PSP, PD and healthy controls. These differences were able to discriminate MSA-P from PSP and confirm that the consequences of cerebellar and pontine pathology known to occur in MSA-P can be detected with DWI during life and help to discriminate it from diseases, which may present with a similar phenotype. None of the rADCs measured, differentiated PSP from healthy controls or from PD, although there was a trend towards an increased rADC in the superior cerebellar peduncle, the frontal white matter and the white

matter of the centrum semi-ovale in PSP, which did not survive a correction for multiple comparisons. PSP could however be differentiated from PD on the basis of the mean (left and right) frontal white matter rADC. The association between rADC values in the brainstem and disease severity supports the hypothesis that tissue damage in these regions contributes to the motor deficit characteristic of PSP and MSA-P. The associations between rADC and brainstem volume supports the established hypothesis that cell damage results in the regional atrophy detected in vivo in volumetric MRI studies.

There was overlap in the individual rADC measurements between the groups, meaning that using rADC values to distinguish PSP and MSA-P has limitations as there will inevitably be false negatives (1 of 11 MSA-P classified as PSP in this study) and false positives (4 PSP patients classified as MSA-P). The rADC cut-off would have to be set at $0.886 \times 10^{-3} mm^2 s^{-1}$ to exclude all PSP patients and this would increase the false negative classification of MSA-P patients.

There was no association between rADC and disease duration in any of the disease groups, however previous studies have identified associations between rADCs in the MCP and Pons, and disease duration in MSA-C (Kanazawa *et al.*, 2004). In MSA-C (cerebellar phenotype), the association between rADC values in the MCP and pons and disease duration may be clearer than in MSA-P where more severe pathology may be expected in the basal ganglia. An alternative explanation is that there was greater variability in the disease duration of the MSA-P subjects we studied (mean, 5.6 years) compared to the MSA-C patients (median, 3 years) (Kanazawa *et al.*, 2004). Disease heterogeneity and measurement error may also influence the observed relationship between disease duration (which is notoriously inexact) and rADC measurements.

Previous studies have demonstrated that rADC measurements in the putamen, globus

pallidus and caudate nucleus discriminated PSP from PD (Seppi *et al.*, 2003). We were not able to reproduce these results.

The patients in our study were of a similar age to those reported previously but both the PSP and MSA-P patient groups in our study had longer mean disease durations. Despite this, the mean rADC values in the putamen were lower in PSP and MSA-P in our study than in those reported previously. The boundaries of the putamen were difficult to clearly identify on our DWI images, both with a b-value of 0 s/mm^2 (which are essentially T2-weighted) and with a b-value of 1000 s/mm^2. This was especially the case in MSA-P where a decrease in T2 signal intensity of the putamen is often present (Bhattacharya *et al.*, 2002). The repeatability and reliability of the rADC measurements was reduced and hence measurement error may have contributed to the lack of group differences in this region.

An association was detected between increased motor deficit measured on the Hoehn and Yahr scale and increased rADC in the upper pons in MSA-P. Motor deficit measured using UPDRS III was associated with rADC in the upper pons and thalamus in PSP. In PSP and MSA-P, the pathological damage to tissue architecture is predominantly present in the brainstem, and it is not surprising that rADCs here correlate with the motor deficit.

The motor thalamus is affected in PSP and a recent study has implicated pathological involvement of this structure in postural instability (Halliday *et al.*, 2005). The thalamic rADC was also associated with FAB score in PSP although frontal white matter rADC values were not. The only previous study to look for associations between frontal white matter ADC values and measures of cognitive function in PSP was negative in five patients (Ohshita *et al.*, 2000). The association between thalamic rADC values and FAB scores, and a lack of association with frontal white matter

rADC, suggests that subcortical structures may influence the subcortical cognitive impairment characteristic of PSP more than cerebral cortical pathology. There is also evidence to suggest that diffusion imaging may be particularly sensitive to damaged thalamic grey matter in CADASIL and that this is strongly associated with sub-cortical cognitive impairment (O'Sullivan *et al.*, 2004).

The association of a lower TIV corrected volume in the pons and cerebellum, with higher rADC values in these regions in MSA-P is to be expected. The pathological hallmark of MSA is alpha synuclein positive glial cytoplasmic inclusions in oligodendroglia. These are widely distributed throughout the brain but are most numerous in the pontine base and cerebellar white matter, resulting in damage in these regions to myelinated structures (Matsuo *et al.*, 1998; Papp and Lantos, 1994). In regions where pathological processes have altered the integrity of the tissue in this way, microstructural barriers to water diffusion are reduced, resulting in an increased ADC.

Another inevitable eventual outcome is regional atrophy due to cell loss. Pathological studies have demonstrated that neuronal cell loss (which is the basis of macroscopic atrophy) is correlated with the severity of glial cytoplasmic inclusions in MSA (Ozawa *et al.*, 2004) supporting the association between TIV corrected pontine volume and the caudal pons ADC in MSA-P. The increased rADC in the MCP in MSA-P may arise as a consequence of direct involvement of oligodendroglia in the MCP but also Wallerian degeneration. This is a potential further explanation for the ADC changes particularly in the cerebellar peduncles, where diffusion changes and T2 hyperintensities have been described following infarcts in the pons (Marx *et al.*, 2004). Both processes would potentially contribute to the association between the ADC in the MCP and cerebellar volume, although the latter process would probably

not damage tissue architecture to the same extent. That a stronger association between midbrain volume and midbrain rADC was not seen in PSP, may be because while the normal cellular architecture is disrupted by the pathological process resulting in cell loss and atrophy, the ADC is not so severely affected because there is less white matter pathology. Pathological involvement of the red nucleus and the substantia nigra in the midbrain may not affect the rADC value but cell loss in these regions does result in atrophy.

The results reported in this chapter are cross sectional and we cannot establish whether rADC changes precede atrophy. However, in AD, diffusion changes in the hippocampus precede volume loss and the full clinical manifestation of the disease (Kantarci *et al.*, 2005), confirming that DWI is sensitive to ultrastructural changes at an early disease stage. Similar longitudinal studies are necessary to establish the sensitivity of ADC measurements in degenerative bradykinetic-rigid syndromes.

Echo-planar DWI has the advantage of rapid acquisition and is therefore less prone to movement artefact, which can affect the clinical utility of other quantitative methods, such as volumetric MRI, in neurodegenerative movement disorders. This makes it an attractive imaging modality in degenerative movement disorders such as PSP.

The rADC values recorded in this chapter for the pons and MCP are similar to those reported by other groups (Kanazawa *et al.,* 2004). The rADC values (pons, MCP, thalamus, caudate nucleus and parietal white matter) previously reported for healthy controls are lower than those in this thesis but this may be explained by the younger age of the control subjects (mean 57.6, SD 12.0 years) in the Japanese study. The frontal rADC values, however, are similar to those reported in previous DWI studies in PSP, although the parietal white matter rADC values are lower (Ohshita *et al.,* 2000). However the patient cohort in the Japanese study was smaller (n = 5) and the

disease duration shorter (3.2 years). This could explain the lower rADC value recorded in the parietal white matter, as this brain region may only be pathologically involved later in the disease course.

9. Serial volumetric imaging analysis

In cross-sectional imaging studies, accumulated tissue losses are unlikely to yield dramatic differences in structural volumes between disease groups, or volumes outside of the normal range in healthy controls because of considerable inter-individual variation in the normal volume of these structures. A shortcoming of cross sectional studies is that they can only *imply* that group differences reflect that atrophy has occurred. In order to calculate the rate of brain atrophy, using serial MRI at two or more time points is necessary.

Serial imaging studies in MSA using qualitative and quantitative area measurements have been undertaken (Horimoto *et al.,* 2000; Horimoto *et al.,* 2002; Konagaya *et al.*, 2002; Watanabe *et al.,* 2002) and registration of longitudinal volumetric MRI brain scans has been thoroughly evaluated in AD (Fox *et al.*, 1996; Fox *et al.*, 1999; Fox and Freeborough, 1997; Jack, Jr. *et al.*, 2000) and FTD (Chan *et al.*, 2001b; Chan *et al.,* 2001a; Schott *et al.*, 2003). However analysis of longitudinal imaging has not been applied to study rates and patterns of atrophy based on volumetric brain measurements in PSP.

Determination of rates of whole brain and regional brain atrophy in PSP compared to MSA-P, PD and healthy controls using registration of serial MRI brain scans in PSP was therefore carried out.

9.1. Serial regional segmentation

9.1.1. Introduction

Whole brain and regional brain atrophy can be assessed, by measuring brain or regional brain volumes at two time points and subtracting the volume at time point two from the same volume measured at time point one. Atrophy rates can then be expressed as a percentage change in brain volume over a specified time interval.

9.1.2. Methods

Patients with PSP, MSA-P and PD who had attended for baseline and follow up MRI scans were included. The digitised structural MR images were transferred to a SUN workstation (Sun Microsystems Inc, Mountainview, CA) and image analysis undertaken using the MIDAS package as outlined in the previous chapters (Freeborough *et al.,* 1997). Post-acquisition correction of heterogeneity artefact was applied as before (Lewis and Fox, 2004). Images were registered to standard space. Serial brain scans were registered to each other (spatially matched using 9 *dof*) applying whole brain to whole brain matching (Woods *et al.,* 1998). Whole brain and regional brain volumes on the baseline and follow up scan were measured using the methodology described in section 6.1. Regional segmentations were undertaken on the baseline and follow up scan at the same time, with the operator blinded to the identity, diagnosis and time sequence of the scans.

For each scan pair, the change in brain and regional volume was calculated by subtracting the volume of the second manually segmented scan region from the volume of the first manually segmented scan region.

The percentage rate of brain or regional atrophy per annum was then calculated by applying the following formula:

$$\text{Atrophy rate (\%)} = (([vol_1 - vol_2] / vol_1) \times 100) \times 365 / \text{Scan interval}$$

Where vol_1 is the brain or regional volume at time point 1, vol_2 the same at time point 2 and scan interval is the time between scan 1 and scan 2 in days.

9.1.3. Results

Twenty-four patients with PSP, eleven with MSA-P, twelve with PD and eighteen healthy controls had an initial scan. Three of the PSP patients were excluded from further follow up as the first scan was of poor quality due to excessive movement artefact. Two patients (one PSP and one MSA-P) were too unwell to attend for a second scan a year later, 2 PD patients withdrew from the study and 1 with MSA-P and 1 with PSP died. One individual classified as PSP at the initial visit developed visual hallucinations and a clinical syndrome more in keeping with a clinical diagnosis of dementia with Lewy bodies (DLB) and was excluded from data analysis. Follow up scans in one patient with PSP and one with PD were of poor quality and were not included. Of the 65 individuals recruited, 53 (17 PSP, 9 MSA, 9 PD and 18 healthy controls), or 82% had serial MRI scans analysed.

The atrophy rates calculated from simple subtraction of segmented images at two time points, are shown in table 9.1. For normally distributed data, a one-way analysis of variance with post hoc Bonferroni testing was used to look for significant differences between group means. For non-parametric data, a Mann-Whitney U test was applied.

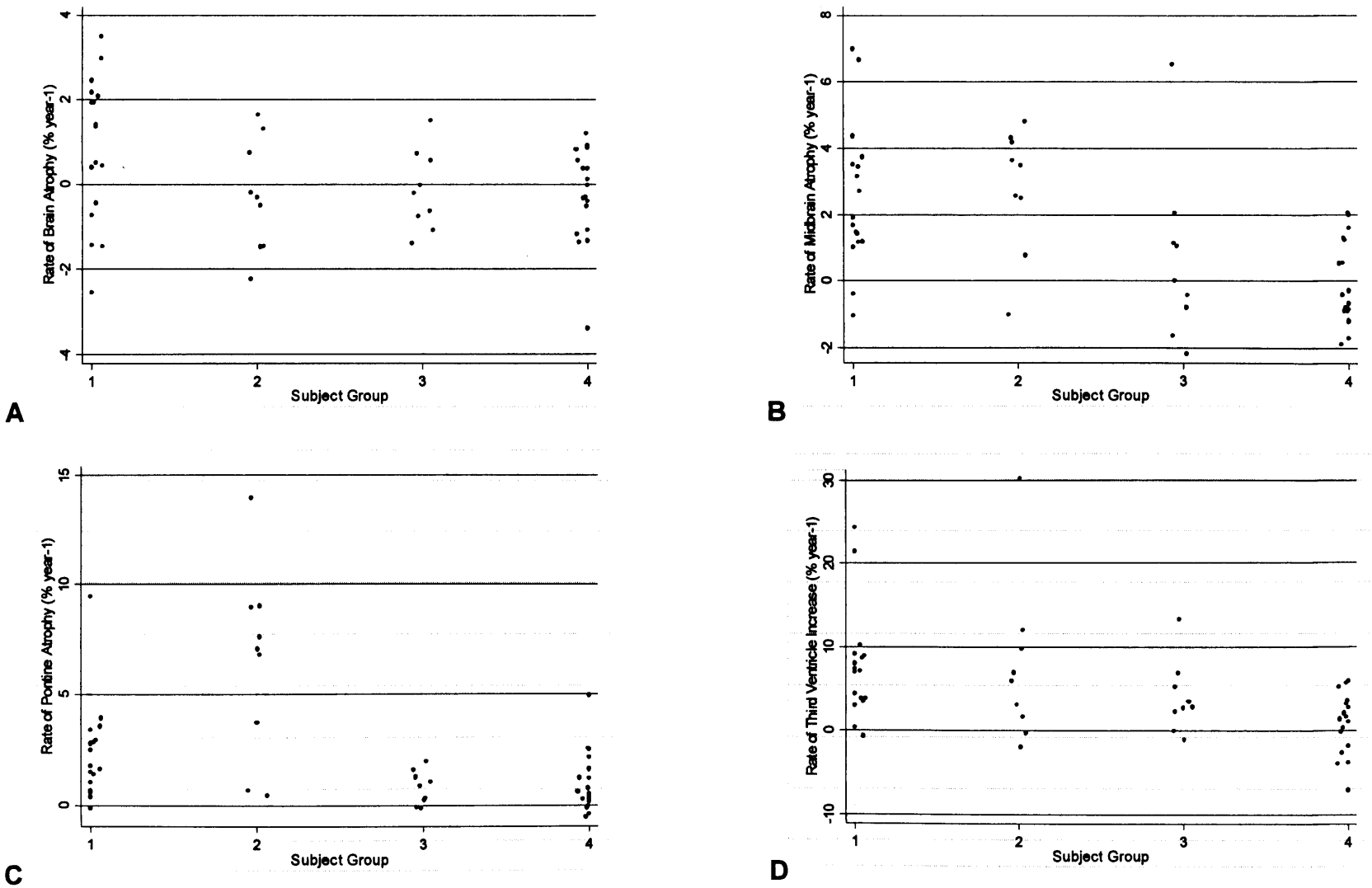
Table 9.1 *Mean (SD) Manually calculated annualised rates of atrophy (% per year)*

	PSP	MSA-P	PD	Healthy Controls	P value
Scan interval (days)	291.6 (58.8)	262.3 (65.5)	268.2 (51.5)	251.9 (39.7)	0.2
Whole brain	0.84 (1.7)	-0.29 (1.4)	-0.14 (0.95)	-0.28 (1.1)	0.07
Midbrain	2.5 (2.2)*	2.8 (1.8)#	0.59 (2.6)	-0.08 (1.2)	0.003
Pons	2.4 (2.2)	6.4 (4.3)	0.82 (0.74)	0.96 (1.4)	*MSA vs. PSP p=0.04* *MSA vs. PD p=0.007* *MSA vs. HC p=0.001* *PSP vs. PD/HC p=0.01*
SCP	3.6 (7.2)	1.4 (7.5)	1.1 (2.3)	1.6 (5.0)	0.7
Lateral Ventricle	9.9 (10.7)	10.0 (11.7)	5.3 (5.7)	3.0 (3.8)	*PSP vs. HC p=0.01*
Third Ventricle	7.5 (6.4)	7.7 (9.8)	4.3 (4.3)	0.88 (3.5)	*PSP vs. HC p<0.001* *MSA vs. HC p=0.04*
Frontal quadrant	0.64 (1.8)	-0.22 (1.3)	-0.18 (1.4)	0.06 (1.5)	0.5
Posterior Inferior	0.44 (2.0)	0.07 (2.1)	-1.1 (2.8)	-0.7 (2.8)	0.4

* PSP vs. PD, p=0.05 vs. HC, p<0.001; # MSA vs. PD, p=0.05, vs. HC, p<0.001. P values in italics are from Mann-Whitney U test for equality of means.

The mean (SD) values give the impression that despite the significant differences between the group mean values, there is considerable overlap in individual atrophy rates calculated in this way. This can be seen in figure 9.1, which shows the manually calculated percentage atrophy rates in the different groups in the brain, the midbrain, and the pons, as well as rates of expansion in the third ventricle.

Figure 9.1 *Scatterplots displaying manually calculated, annualised atrophy rates (%)*



A- whole brain, B- Midbrain, C-Pontine, D- Third Ventricle. 1=PSP, 2=MSA-P, 3=PD, 4=Healthy controls

9.1.4. Conclusions

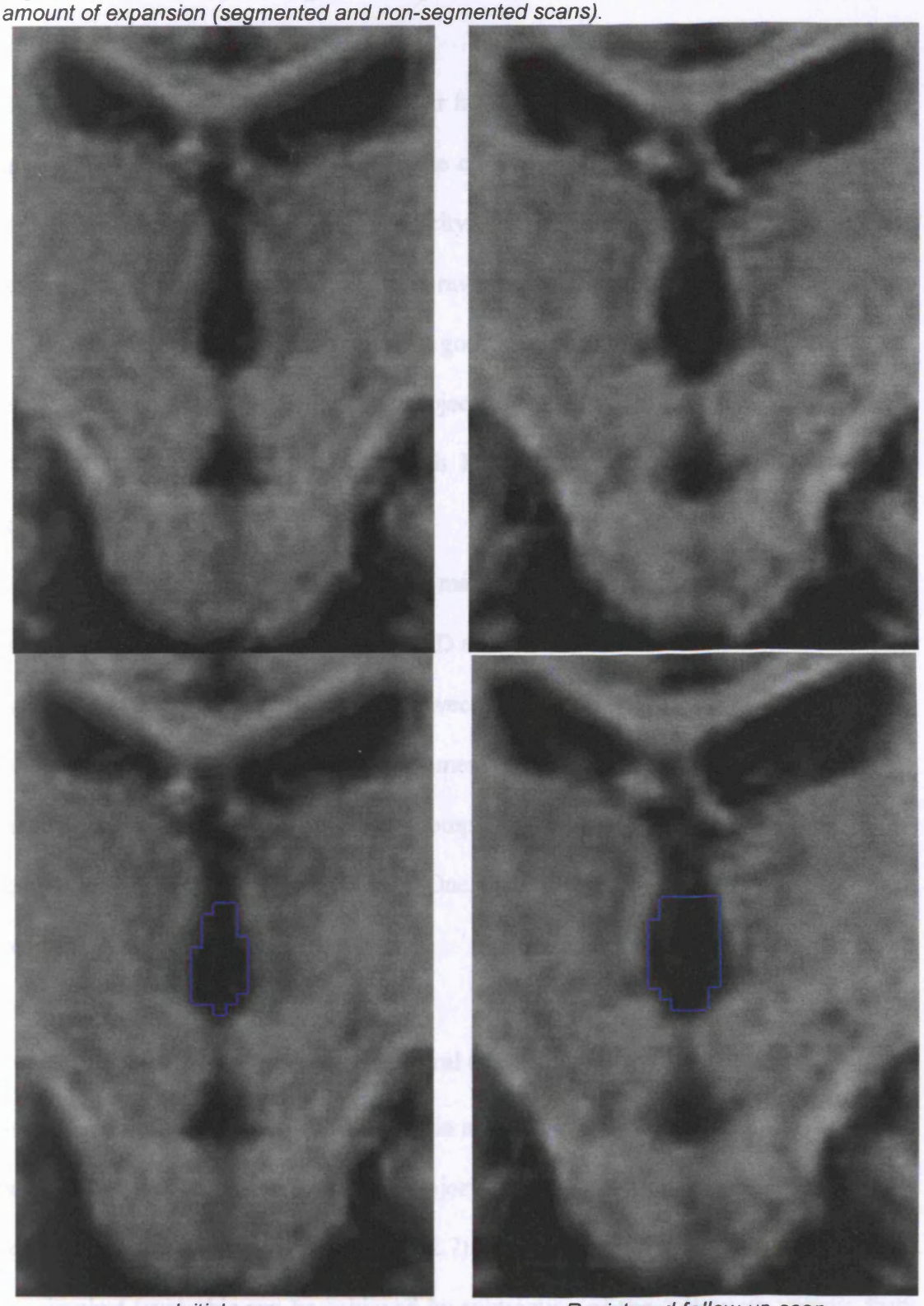
Manually calculated whole brain atrophy rates were not significantly different between the groups ($p=0.07$). While the mean whole brain atrophy rate in the PSP group was greater than in MSA-P, PD and healthy controls, the large inter-individual variation in atrophy rates calculated in this way results in considerable overlap between the groups.

Significant differences between mean midbrain, pontine, lateral ventricle, and third ventricle rates of atrophy or expansion were detected between the groups. Surprisingly, midbrain atrophy rates in MSA-P were similar to those in PSP. Pontine atrophy rates were greatest in MSA-P, but in PSP, atrophy rates in these regions were also significantly greater than in PD and healthy controls.

One subject with PD had a rate of midbrain atrophy well above the rest of the values (figure 9.1). A single individual with MSA-P had a rate of third ventricle expansion well above the values of the rest of the group. The MRI scans in these patients were visually checked for obvious segmentation errors, but no clear erroneous voxels were included in the region concerned. Figure 9.2 shows that the third ventricular enlargement in the MSA-P case was indeed marked and that segmentation of the region was accurate.

Despite this, manual measurements are likely to be more prone to measurement error than semi or fully automated measurements and this measurement error is likely to be one factor contributing to the variance in individual atrophy rates within the groups. True heterogeneity of atrophy rates within groups is another important factor with variability in whole brain atrophy rates in healthy controls being less marked in PSP.

Figure 9.2 *Third ventricle enlargement. Registered MRI scans in MSA demonstrating large amount of expansion (segmented and non-segmented scans).*



Initial scan *Registered follow up scan*

Although the median disease duration from onset to death in PSP is approximately 6 years (Litvan *et al.*, 1996c), the disease is known to have a variable rate of progression (Williams *et al.*, 2005).

Misclassification of individuals is a further factor that may lead to variation in atrophy rates within a disease group, however the clinical diagnostic criteria applied in this study have good sensitivity and specificity, and all individuals were followed up, beyond the time of the second scan and reviewed to ensure accuracy of the clinical diagnosis, at a centre in which there is a good record of clinical diagnostic accuracy (Hughes *et al.*, 2002). To date, of the subjects included in this serial imaging section of the study, 2 with PSP and one with MSA-P have had the clinical diagnosis confirmed at post mortem.

Brainstem atrophy rates calculated using manual measurements on serial MRI scans, were greater in PSP and MSA-P than in PD and healthy controls. Whole brain atrophy rates reveal no significant differences between the groups.

One way of assessing whether measurement error has a marked influence on the manually calculated atrophy rates, is to compare them with atrophy rates derived from semi-automated measurement methods. One such method is the brain boundary shift integral (BBSI).

9.2. The brain boundary shift integral (BBSI) and rigid body registration

Several methods of measuring whole brain atrophy from serially acquired volumetric brain scans have been developed. The majority rely on positional matching of serially acquired scans (registration, see chapter 2.7), followed by a quantification process. At the simplest level, this can be achieved by subtraction of the follow up volume from the baseline volume. This requires that whole brain and regional brain volumes are

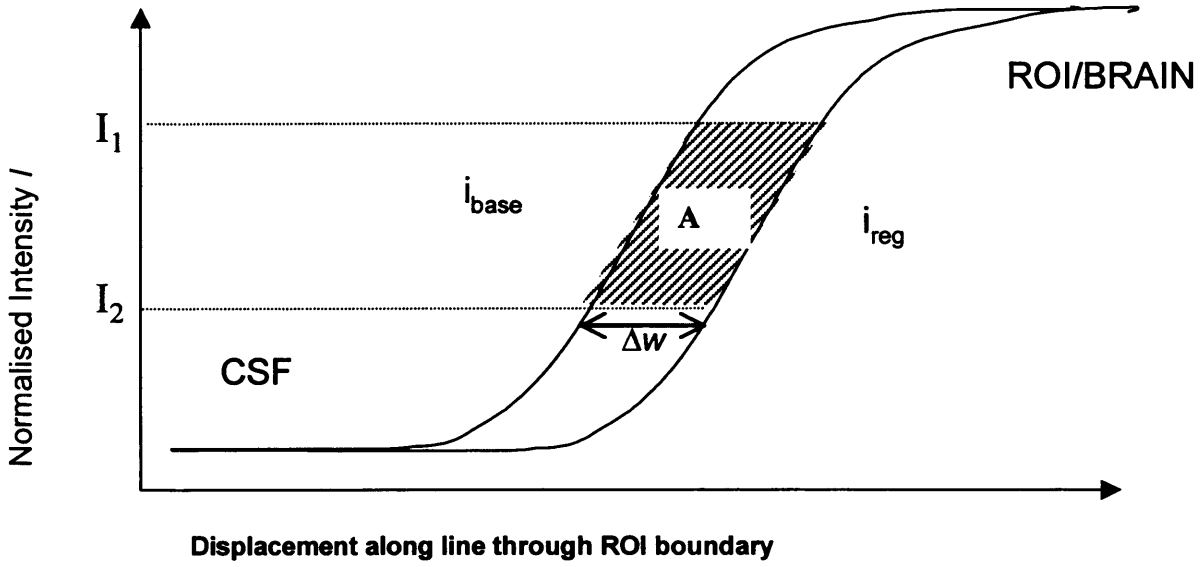
corrected for total intracranial volume (TIV), or are registered using a 9 *dof* registration to provide protection against voxel drifts. These subtraction methods are more dependent on individual measurements at each scan time point and as such are likely to result in more measurement error.

9.2.1. Introduction and theoretical aspects

The brain boundary shift integral (BBSI) has been proposed as a measure of cerebral volume change from serial MRI (Freeborough and Fox, 1997; Gunter *et al.*, 2003). This technique is based on identifying the brain tissue in two different scans from the same individual, registering the two scans spatially, and then subtracting scan intensities in the area between the borders of the aligned brains. Once the two scans have been segmented, the remainder of the process is fully automated, requiring only two user-defined parameters that define the location and the width of the sampling (image intensity) window (see figure 9.3). The BBSI increases the precision of estimation of atrophy (as compared to the differences in manually segmented volumes) by eliminating some random measurement error. However, it also introduces a systematic underestimate of volume change due to some of the changes in scan intensity not spanning the parameter-defined window.

The BBSI operates on the following principles. Viewed in cross section, the boundary between brain and CSF, on a T1 weighted MRI scan, can be seen as an intensity curve crossing from brain (high intensity to CSF (low intensity). This intensity curve can be considered for the baseline $i_{base(x)}$ and the registered follow up scan $i_{reg(x)}$. If the boundary in the second scan has shifted by a certain amount (Δw) then this can be plotted schematically (figure 9.3).

Figure 9.3 ***The brain boundary shift integral (BBSI)***



A representation of an intensity profile through a boundary of brain on a baseline (i_{base}) and a registered repeat (i_{reg}) MRI scan. During the interval, the boundary has shifted by an amount Δw, which we determine as the area A (divided by I_1-I_2).

The area labelled A in figure 9.3 represents the integral with respect to the intensity of the difference in position of the profiles over the interval $[I_1, I_2]$; hence

$$A = (I_1 - I_2)\,\Delta w$$

The area A may alternatively be calculated as integral of the intensity profile of the registered scan from that of the baseline over the range I_1-I_2.

Thus Δw, the boundary shift may be calculated and extended to three dimensions, in order to obtain the *brain boundary shift integral* (BBSI). This represents the total volume traversed by the boundaries of the brain going from the first scan to the second scan, and is a direct measure of change in volume. Calculation of the BBSI requires that appropriate values are selected for I_1 and I_2; both should be within the

range of the intensity transitions at the boundaries of the brain. Using I_1 and I_2, a window size can be calclated I_w as can the window centre Ic.

9.2.2. BBSI and regional BSI Methods

As in the case of the manually measured regions of interest, serial brain scans were first registered (spatially matched using 9 *dof*) applying whole brain to whole brain matching (Woods *et al.*, 1998). The percentage annual brain volume loss was calculated from the BBSI, taking the BBSI as the volume (ml) of brain tissue loss and applying the following equation:

$$\text{Atrophy rate (\%)} = ((\text{BBSI} / \text{vol}_1) \times 100) \times 365 / \text{Scan interval}$$

Local BSIs were calculated by restricting the BSI calculation to the defined region of interest. Serial brain images were registered applying whole brain to whole brain matching as above. ROI were then further spatially matched by registering baseline ROI to the follow up ROI using a 6 *dof* registration. The region created by subtracting an eroded (by one voxel) intersection region from a dilated (by one voxel) union region of the resulting two brain regions was generated and a dilated (by one voxel) baseline and follow up ROI used to mask this in order to allow calculation of the BSI in this anatomical area of interest. This is illustrated schematically in figure 9.4. For the cerebellum, only the baseline cerebellar volume was manually segmented and the regional BSI calculated using only this baseline (dilated by one voxel) region. This technique has been applied to hippocampal ROI measurements in Alzheimer's disease (Barnes *et al.*, 2004). All registrations were visually checked for accuracy and data excluded if the registration of the local region was considered to be poor as a consequence of movement artefact.

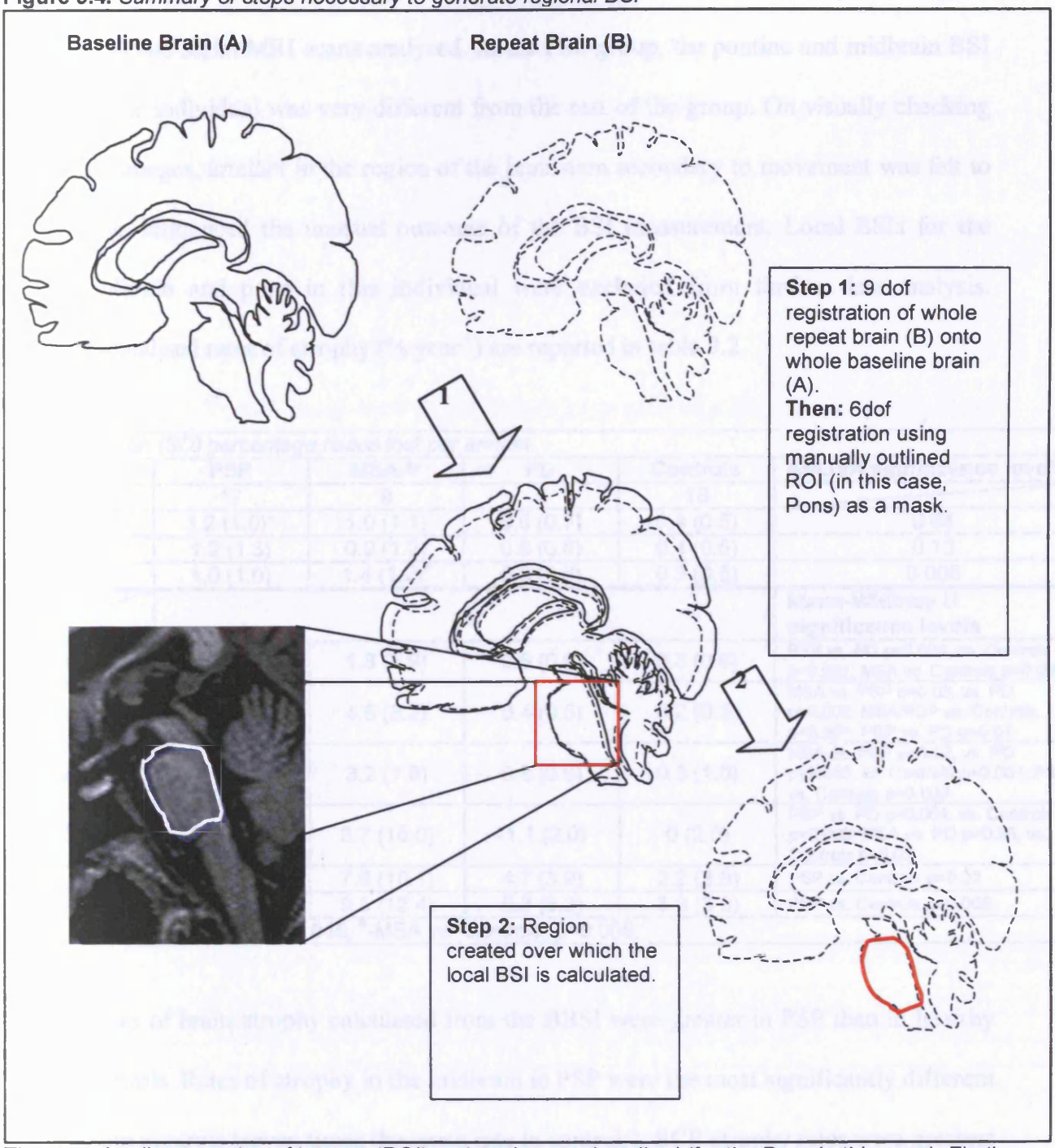
Annualised changes were calculated by taking the BBSI and the regional BSIs in each group (ml) and adjusting for scan interval. Atrophy rates were calculated as an annual percentage change from the baseline volume.

Total intracranial volume was not used as a covariate in atrophy rate calculations as regression analysis reported no relationship between TIV and atrophy rates. A one-way analysis of variance test was used to compare whole brain and regional atrophy rates for normally distributed data, and a Mann-Whitney U test was used to compare atrophy rates for non-parametric data.

Pitman's test was used to assess whether BBSI and regional BSI's reduced the variance in atrophy rates reported using the manually calculated rates of atrophy.

Logistic regression models were applied to assess the ability of BSI derived atrophy rates to discriminate between PSP and MSA-P.

Figure 9.4. *Summary of steps necessary to generate regional BSI*



The steps leading to region of interest (ROI) registration are described (for the Pons in this case). Firstly, a nine degrees of freedom (9dof) registration is used to match whole brains, then a 6dof registration is applied to match the ROI. Following this, the region that contains the ROI boundary shift is generated and the ROI BSI calculated in this region only.

9.2.3. Results

Of 65 individuals recruited, 53 (17 PSP, 9 MSA, 9 PD and 18 healthy controls), or 82% had serial MRI scans analysed. In the PSP group, the pontine and midbrain BSI in one individual was very different from the rest of the group. On visually checking the images, artefact in the region of the brainstem secondary to movement was felt to have influenced the unusual outcome of the BSI measurement. Local BSIs for the midbrain and pons in this individual were excluded from further data analysis. Annualised rates of atrophy (% year^{-1}) are reported in table 9.2.

Table 9.2. *Mean (SD) percentage tissue lost per annum*

	PSP	MSA-P	PD	Controls	ANOVA significance levels
N	17	9	9	18	
Brain	1.2 (1.0)*	1.0 (1.1)	0.6 (0.7)	0.4 (0.5)	0.04
Frontal	1.2 (1.3)	0.9 (1.2)	0.6 (0.8)	0.4 (0.5)	0.13
PI	1.0 (1.0)	1.4 (1.2)#	0.8 (0.6)	0.3 (0.5)	0.008
					Mann-Whitney U significance levels
Midbrain	2.2 (1.5)	1.8 (1.9)	0.6 (0.9)	0.3 (0.6)	PSP vs. PD p=0.004, vs. Controls p<0.001; MSA vs. Controls p=0.006
Pons	1.5 (1.3)	4.5 (3.2)	0.4 (0.5)	0.2 (0.3)	MSA vs. PSP p=0.03, vs. PD p=0.002; MSA/PSP vs. Controls p<0.001; PSP vs. PD p=0.01
Cerebellum	1.4 (1.1)	3.2 (1.9)	0.8 (0.6)	0.3 (1.0)	MSA vs. PSP p=0.02, vs. PD p=0.005, vs. Controls p<0.001; PSP vs. Controls p=0.009
SCP	3.5 (4.0)	3.7 (10.0)	-1.1 (2.0)	0 (2.5)	PSP vs. PD p=0.004, vs. Controls p=0.008; MSA vs. PD p=0.05, vs. Controls p=0.03
Third Vent.	6.1 (6.0)	7.6 (10.1)	4.7 (3.9)	2.2 (3.8)	PSP vs. Controls p=0.02
Lateral Vent.	9.6 (9.8)	9.5 (12.4)	5.3 (5.2)	3.3 (3.3)	PSP vs. Controls p=0.008

ANOVA *-Controls vs. PSP, p=0.045; #-MSA vs. Controls, p=0.008.

Rates of brain atrophy calculated from the BBSI were greater in PSP than in healthy controls. Rates of atrophy in the midbrain in PSP were the most significantly different from controls (seven times the mean rate in controls). SCP atrophy rates were greatest in PSP, but the variability in the measured atrophy rate limited the significance of the group differences. Pontine atrophy rates were greatest in MSA-P (20 times the mean rate in controls and three times the mean rate in PSP).

Pitman's test confirmed that using the BSI to calculate whole brain and regional atrophy rates, rather than the difference in manually measured volumes, significantly reduced the variability in the atrophy rates reported for whole brain ($p<0.001$), midbrain ($p=0.001$), pons ($p=0.001$) total frontal ($p=0.002$) and PI ($p<0.001$) regions, but not for the SCP, the lateral or third ventricles. This reduced variability in atrophy rates can be appreciated in figure 9.5, which shows individual BSI derived and manually derived atrophy rates for the whole brain, the pons, the midbrain and percentage expansion rates for the third ventricle.

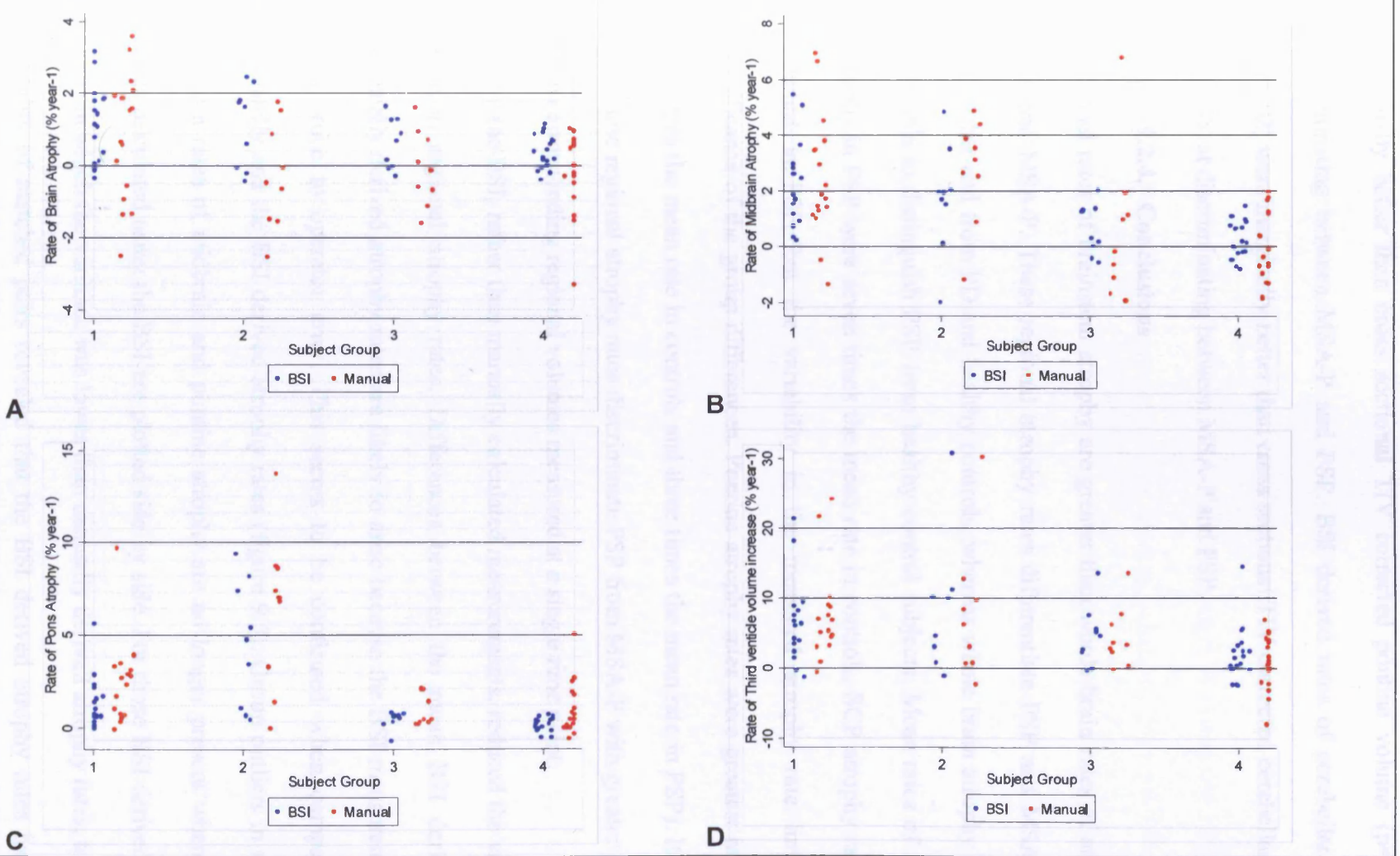
In addition to calculating rates of ventricular expansion as a percentage change from the baseline measurement, rates were also calculated as a percentage of baseline whole brain volume. This second method was employed as for individuals with a small baseline ventricular volume, a small change results in a large percentage expansion over one year. For those with a larger baseline volume, the same volume change over one year will be much smaller when expressed as a percentage. The results of ventricular rates of expansion calculated in this way are presented in table 9.3.

Table 9.3. *Mean (SD) ventricular expansion per annum(% of whole brain volume)*

	PSP	MSA-P	PD	Controls	ANOVA significance levels
N	17	9	9	18	
Third Vent.	0.02 (0.02)	0.009 (0.001)	0.007 (0.006)	0.003 (0.004)	PSP vs. Controls p=0.001
Lateral Vent.	0.35 (0.38)	0.21 (0.24)	0.19 (0.20)	0.08 (0.08)	PSP vs. Controls p=0.001

While the ventricular expansion expressed as a percentage of brain volume results in a different value to that seen when expressed as a percentage of baseline ventricular volume, the rates of expansion are still only significantly different in PSP compared with healthy controls.

Figure 9.5 *Scatterplots comparing BSI derived and manually calculated, annualised atrophy rates (%)*



A- whole brain, B- Midbrain, C-Pontine, D- Third Ventricle. 1=PSP, 2=MSA-P, 3=PD, 4=Healthy controls

From logistic regression models, rates of BSI derived pontine atrophy (p=0.02) were marginally better than cross sectional TIV corrected pontine volume (p=0.04) at discriminating between MSA-P and PSP. BSI derived rates of cerebellar atrophy (p=0.02) were marginally better than cross sectional TIV corrected cerebellar volume (p=0.05) at discriminating between MSA-P and PSP.

9.2.4. Conclusions

Regional rates of brainstem atrophy are greater than whole brain rates of atrophy in PSP and MSA-P. These regional atrophy rates differentiate PSP and MSA-P from each other and from PD and healthy controls, whereas whole brain atrophy rates are only able to distinguish PSP from healthy control subjects. Mean rates of midbrain atrophy in PSP were seven times the mean rate in controls. SCP atrophy rates were greatest in PSP, but the variability in the measured atrophy rate limited the significance of the group differences. Pontine atrophy rates were greatest in MSA-P (20 times the mean rate in controls and three times the mean rate in PSP). It is likely that these regional atrophy rates discriminate PSP from MSA-P with greater accuracy than corresponding regional volumes measured at a single time point.

Using the BSI, rather than manually calculated measurements, reduced the variability seen in regional atrophy rates. Differences between the mean BSI derived and manually derived atrophy rates are likely to arise because the BSI measurements are less prone to operator error. This seems to be confirmed when comparing the manually and the BSI derived atrophy rates (figure 9.5). Group outliers in manually derived rates of midbrain and pontine atrophy are no longer present when atrophy rates calculated using the BSI are plotted side by side. For those BSI derived atrophy rates in which the variance was lower than manually derived atrophy rates, testing for equality of matched pairs revealed that the BSI derived atrophy rates had a bias

towards a greater rate of brain, frontal quadrant and PI atrophy ($p<0.001$) and a lower rate of pontine atrophy ($p<0.001$).

Serial volumetric imaging studies have so far been confined to single case reports or individual cases within the context of cross sectional studies (Paviour *et al.*, 2004b; Schott *et al.*, 2003). A previous study of MRI changes over time in MSA relied largely on qualitative interpretation of imaging changes, but provided useful information regarding the development of certain imaging features over time (Horimoto *et al.*, 2002). Other studies have utilised area measurements to determine atrophy rates in MSA (Horimoto *et al.*, 2000; Konagaya *et al.*, 2002; Watanabe *et al.*, 2002). Konagaya et al studied nine MSA patients in whom the mean scan interval was seven years. The ratio of cerebral hemisphere area to intracranial area decreased from a mean of 82% to 70% suggesting a hemisphere atrophy rate of 2% per annum. Brain atrophy rates reported in this thesis are of a lower magnitude. However area measurements cannot be directly extrapolated to volume change due to the non-uniform distribution of brain tissue loss, as demonstrated in this study. As such, whole brain atrophy rates based on area measurements may give different rates depending upon at which level, the area is measured, although the true rate of whole brain atrophy is the same. In fact, longitudinal imaging studies based on area measurements in MSA, support the presence of non-uniform atrophy, and conclude that cerebellar and pontine atrophy are more significant in MSA with predominant cerebellar signs (Watanabe *et al.*, 2002). These regions had the greatest rates of atrophy in the MSA-P patients in our study, however our patients had a Parkinsonian phenotype. This confirms that atrophy in these regions is still significant when Parkinsonism is the predominant clinical feature. This finding is supported by a large pathological study

which concluded that pontine pathology is present in both MSA-P and MSA-C (Ozawa *et al.,* 2004).

The BSI is semi-automated and because it calculates change directly from differences between MR images, is less vulnerable to errors of manual outlining. It has also been shown to be a repeatable and reliable measure of whole brain atrophy in AD (Fox and Freeborough, 1997). The results of this study suggest that in these neurodegenerative diseases, in which the pathological load has a predilection for brainstem structures, regional rather than whole brain atrophy rates are better markers of disease progression. Case reports of MSA and PSP with pathological correlation of serial imaging studies in life suggest that the regional rates of atrophy that distinguish these conditions are associated with greater pathological load (Paviour *et al.,* 2004b; Schott *et al.,* 2003). The study of a larger number of patients in this thesis confirms these findings.

The rates of regional atrophy within the subject groups reported in this study are heterogeneous. This can partly be explained by manual segmentation errors or scan artefact, but a significant component is likely to be due to true differences in rates of disease progression between individual patients. Although one study based on area measurements, found only weak associations between rates of atrophy and measures of disease severity suggesting that atrophy rates may be relatively constant throughout the disease (Konagaya *et al.,* 2002).

Image analysis methods cannot completely control for differences in scan acquisition and errors may arise due to registration difficulty because of movement artefact or quality differences between the scans. Inevitably, with all of these regional measurements conducted in this thesis, there are some arbitrary decisions about which

structures to include. Errors related to these measurements as well as acquisition may occur and important changes may in fact be under-estimated.

Finally, atrophy rates in these diseases may well be increased relative to healthy controls in the earliest stages of the illness. MRI studies at this stage with a repeat scan (e.g 6-12 months later) may be warranted. It is also at this very early stage that a disease-modifying drug may be at its most useful and so early identification is likely to become increasingly important.

10. Fluid registration

10.1. Introduction

As demonstrated in chapters 7 and 9, accurate positional matching (registration) and subtraction of 3D MR brain scans allows detection of volume differences in an individual brain over time.

The image registration described so far has been achieved using a rigid body transform allowing detection of positional shifts at high contrast boundaries. Positional shifts in uniform tissue regions are also likely to occur but do not result in differences on the subtraction image. Hence the exact location of any tissue changes that bring about the boundary shifts are not revealed.

Large changes in brain tissue volume mean that some of the boundary shifts that occur, reflect changes in the relative positions of structures rather than tissue loss simply at the nearest brain/CSF tissue boundary.

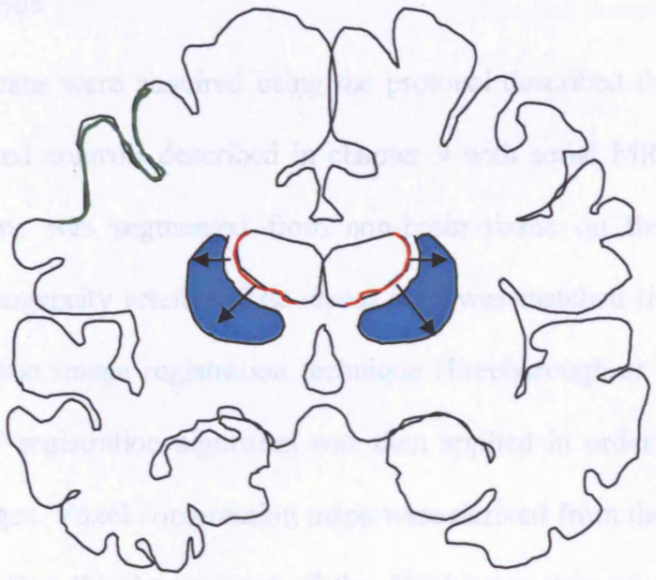


Figure 10.1 *Schematic representation of fluid vs. rigid body registration.*
In rigid body registration, the results are represented as a shift at tissue boundaries (BSI) where red represents expansion and green a region of contraction. The brain is not a rigid body however and so ventricular expansion may occur because of tissue loss distant from the brain/CSF boundary (blue region). A non-linear registration overcomes some of these problems and allows regions of tissue volume loss within the brains structure to be identified.

A non-linear or "fluid" registration overcomes some of these problems (Freeborough and Fox, 1998). Some of the principles behind this method of image analysis have been discussed in chapter 2.7.3 and this technique was used to study the PSP case and the MSA case discussed in chapter 3.5 (Paviour *et al.,* 2004b; Schott *et al.,* 2003). The method is dependent on being able to determine the general deformation that brings one scan into exact structural correspondence with another.

MRI scans from the same patient possess extremely similar topology. "Fluid" registration maintains this topology through the deformation process that registers a follow up scan to a baseline scan. Image intensities on the baseline (Bx) and on the repeat (Rx) scan are denoted. For each voxel position (x), a displacement vector U(x) is obtained. Determining the deformation field Ux that yields point-by-point structural correspondence between the baseline and the repeat scan represents the registration, which is driven to minimise image mismatch.

10.2. Methods

T1 weighted MRI scans were acquired using the protocol described throughout the study. All patients and controls described in chapter 9 with serial MRI scans were assessed. Brain tissue was segmented from non-brain tissue on the scans after correction for inhomogeneity artefact. The repeat scan was matched to the baseline scan with an automatic image registration technique (Freeborough *et al.,* 1996). A non-linear or "fluid" registration algorithm was then applied in order to determine local structural changes. Voxel compression maps were derived from the deformation field Ux, by calculating the determinant of the Jacobian matrix at each point to provide an estimate of volumetric change. These maps are displayed as a colour image overlaying the volumetric MRI scan to allow association of patterns of change

with neuro-anatomically significant regions. This also allows the pattern of regional atrophy occurring in an individual to be demonstrated clearly.

10.3. Results

The fluid registration of an individuals serial scans currently requires several hours of SUN workstation processing time, but can be set to run automatically without the need for an operator to be present. This processing time is rapidly reducing with improving processor speeds.

Individual examples of the subtraction images are shown in figures 10.2-10.7. In addition to the volume loss described in chapter 9, the brains of individuals with PSP and MSA-P clearly undergo deformation and positional shifts as brain atrophy progresses.

Derived voxel compression maps were different between the patients with PSP and those with MSA-P and PD. Consistent patterns of tissue contraction were demonstrated in PSP in the midbrain, the frontal cortical areas and frontal and parietal white matter in PSP (figures 10.2, 3 and 4). In MSA-P much more dramatic cerebellar tissue contraction was evident with consistent disproportionately high rates of tissue contraction in the middle cerebellar peduncle. Compared to PSP, the superior cerebellar peduncle was unaffected in MSA-P and this is demonstrated in figure 10.3 compared to 10.5. Patients with MSA-P also had evidence of increased rates of cortical and white matter tissue contraction (figure 10.6), particularly around the region of the basal ganglia structures. Clear sparing of the midbrain with disproportionate atrophy of the pons was also evident in the MSA-P cases (figure 10.6).

PD and healthy control subjects demonstrated little in the way of tissue contraction.

Figure 10.2 MRI with voxel compression map and colour overlay in PSP.
Tissue contraction (green) is evident, particularly in the midbrain and the brainstem, the collicular plate and thalamus and in the cerebellum.

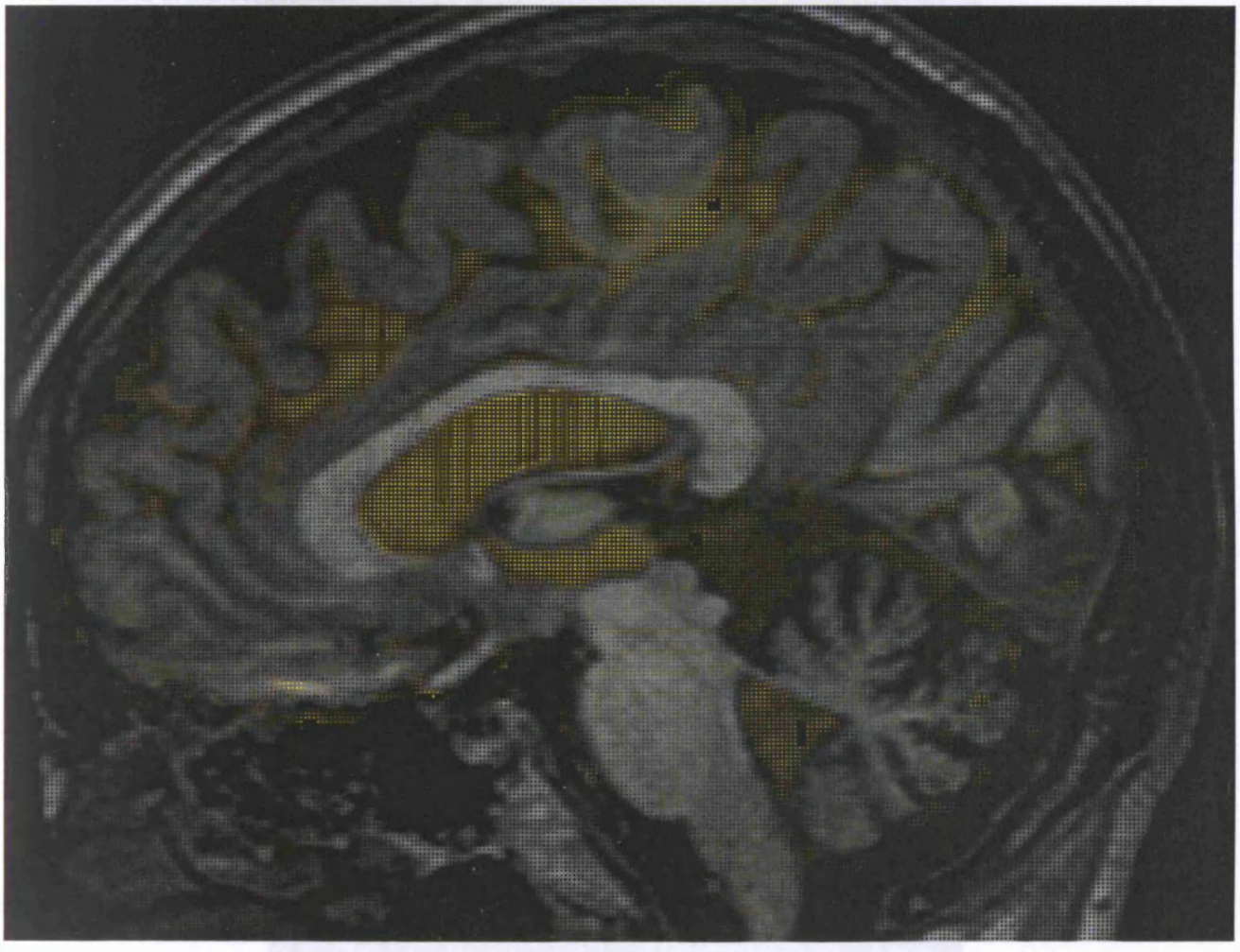


Figure 10.3 MRI with voxel compression map and colour overlay in PSP.
Increased rates of tissue loss (green) in the superior cerebellar peduncle, with relative sparing of the pons is evident in the coronal and axial image. The frontal and parietal lobes also demonstrate atrophy. In the coronal scan (top), expansion of the ventricles is evident.

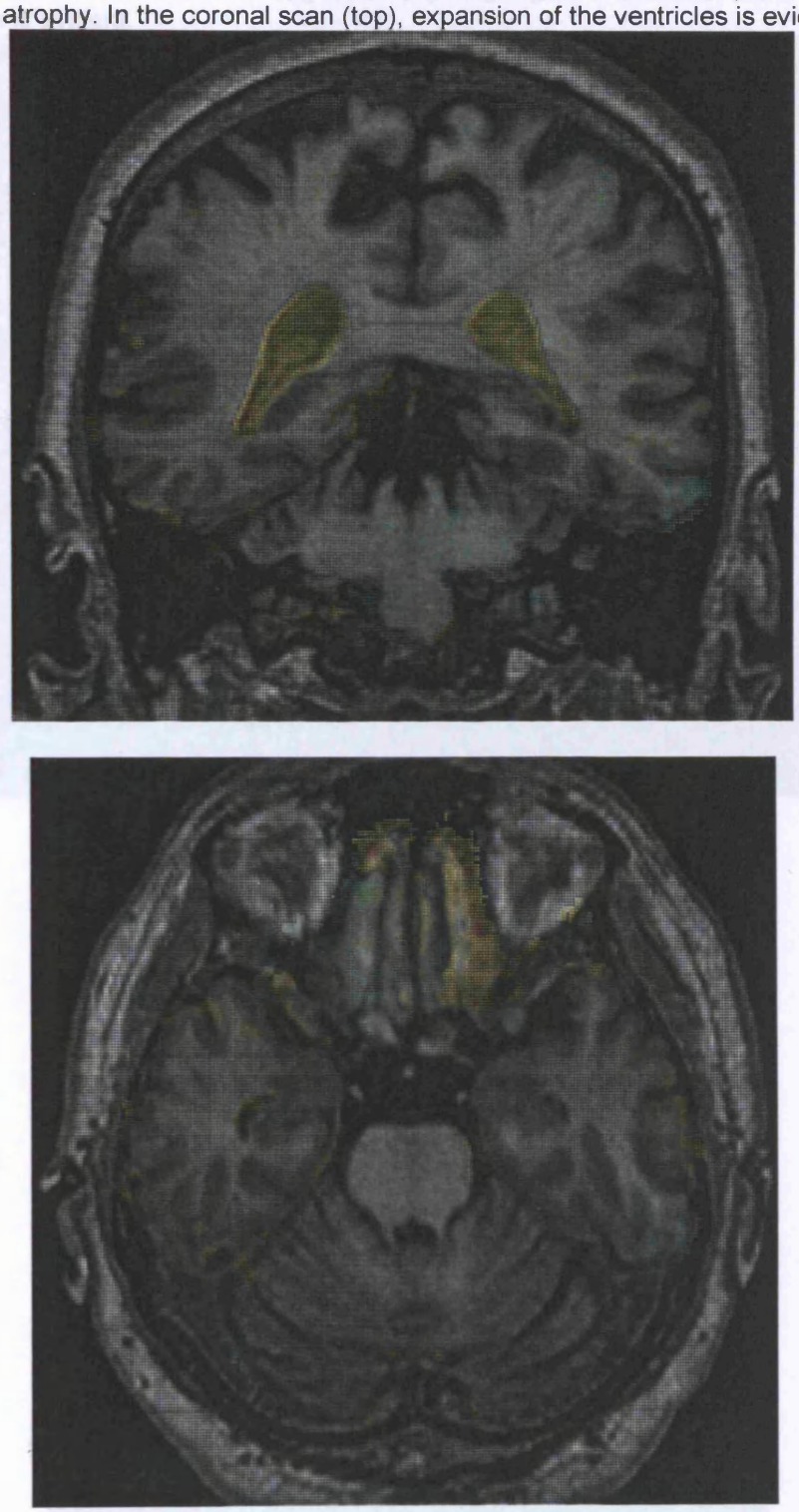


Figure 10.4 MRI with voxel compression map and colour overlay in PSP.
Increased rates of tissue loss (green) in the frontal lobe with relative sparing of the posterior cortical structures is evident in the sagittal and axial images. In the axial scan, expansion of the ventricles is evident.

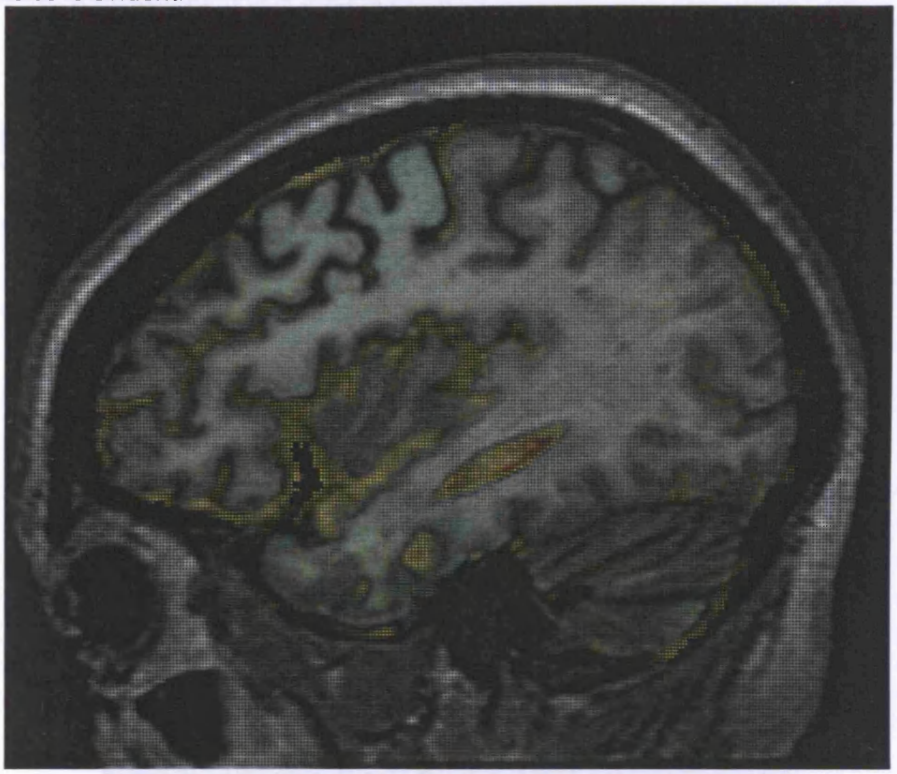


Figure 10.5 MRI with voxel compression map and colour overlay in MSA-P.
An Increased rate of tissue loss (green) in the pons with relative sparing of the superior cerebellar peduncle is evident in the image. In the coronal scan, disproportional involvement of the middle cerebellar peduncle is clear.

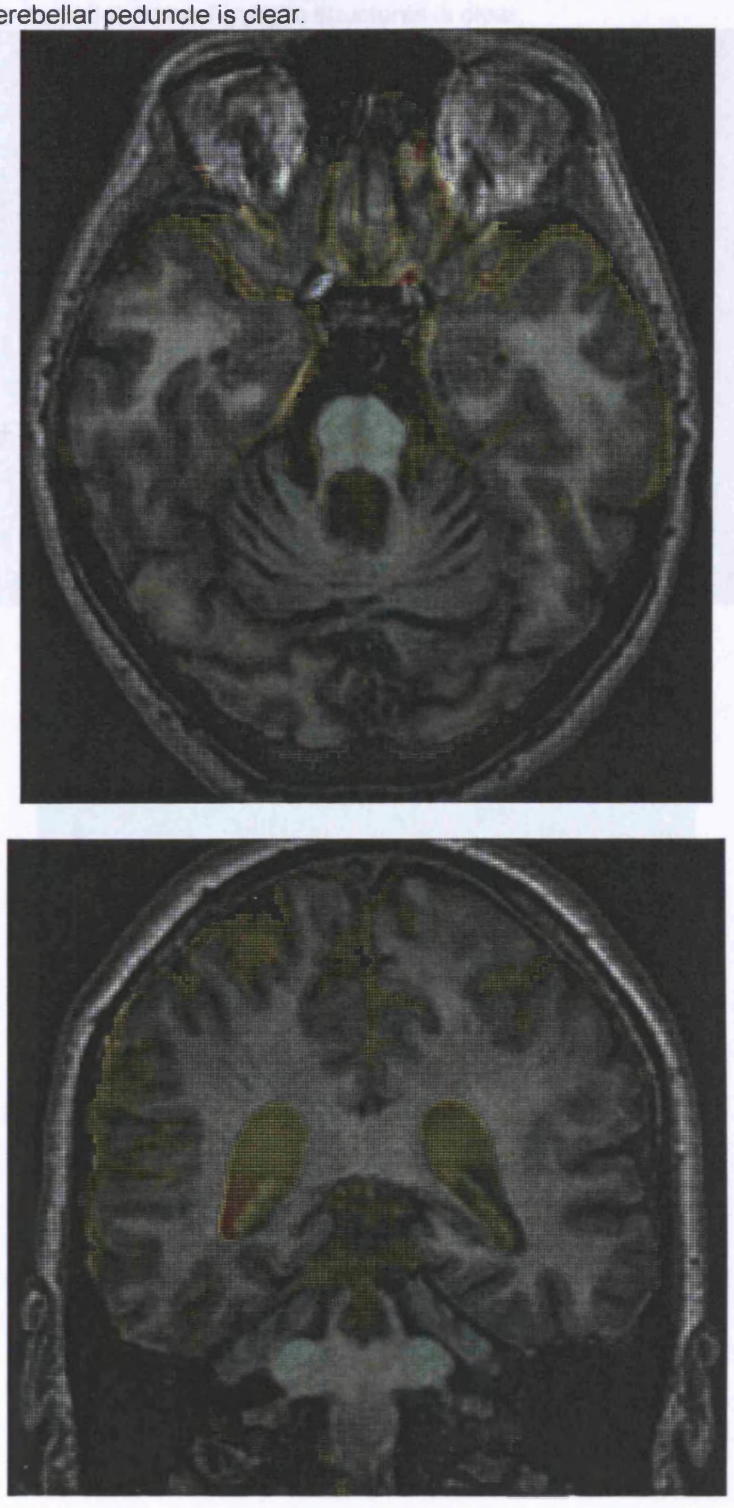
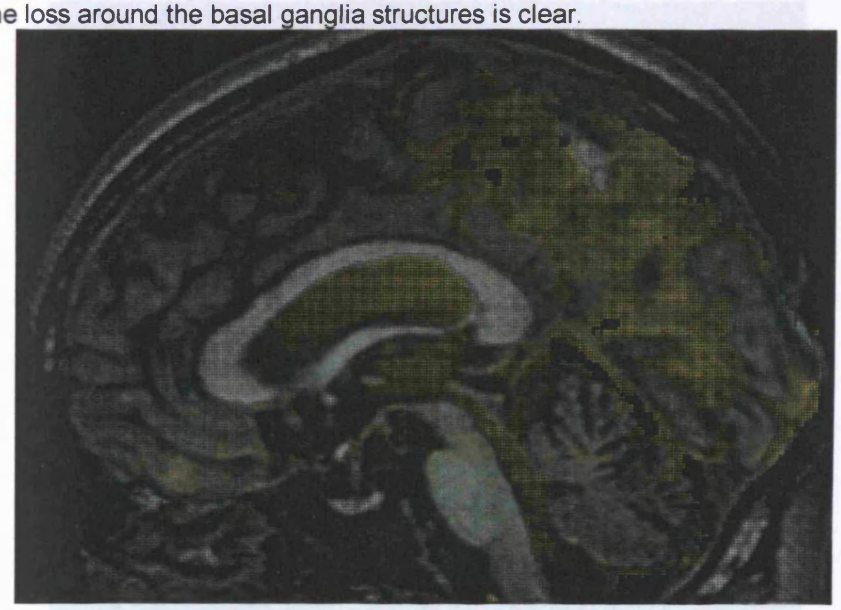


Figure 10.6 MRI with voxel compression map and colour overlay in MSA-P.
An increased rate of tissue loss (green) in the pons with relative sparing of the midbrain is evident in the sagittal image. In the axial scan, diffuse atrophy in the white matter is evident and volume loss around the basal ganglia structures is clear.



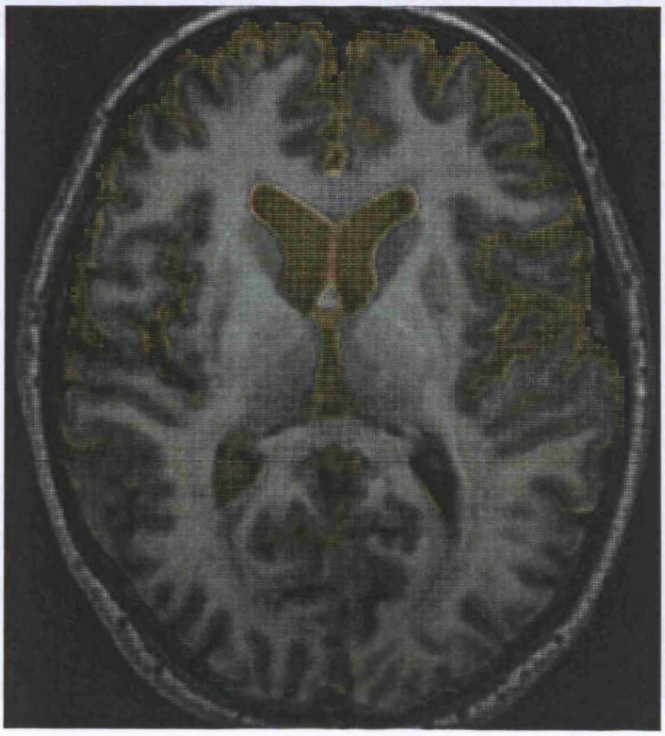
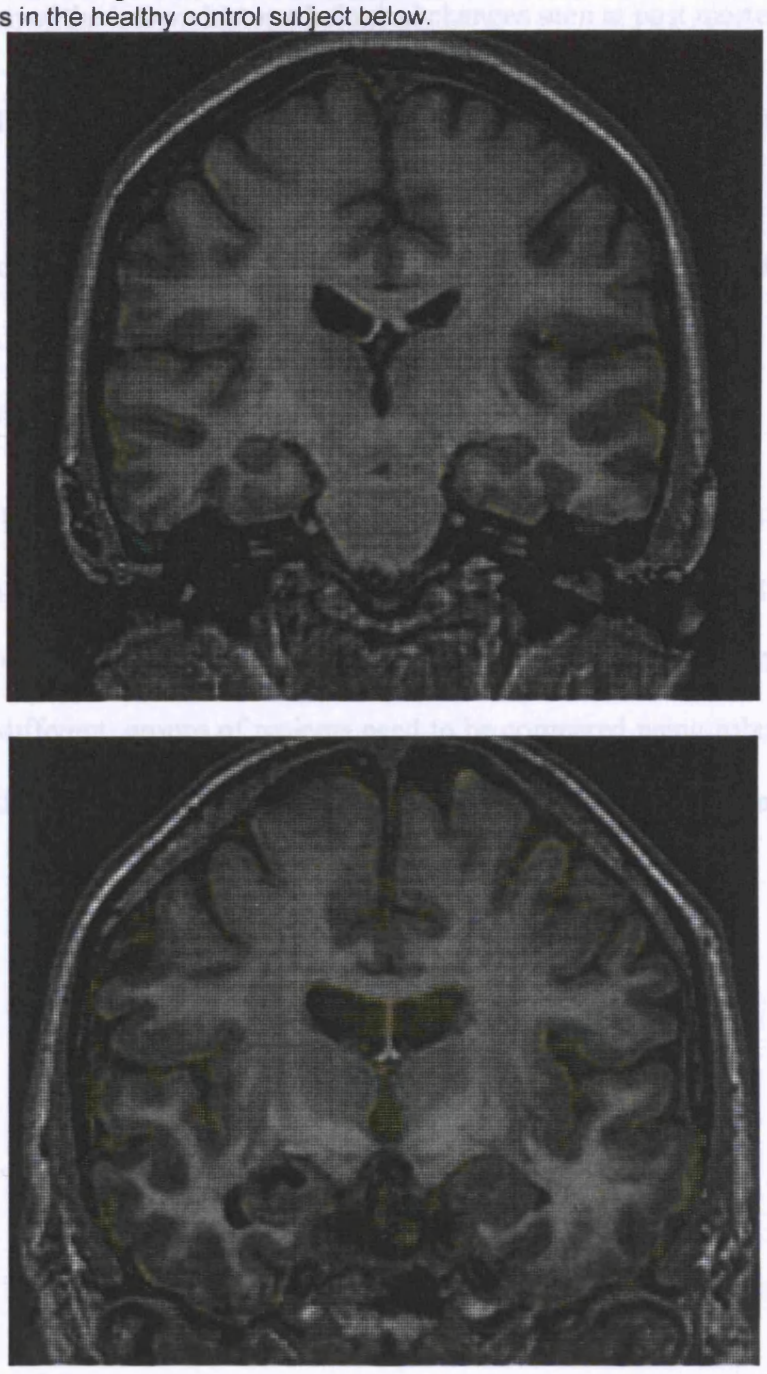


Figure 10.7 MRI with voxel compression map and colour overlay in PD and a healthy control. There is minimal change between the baseline and follow up scans, most notably in the temporal lobes in the healthy control subject below.



10.4. Conclusions

The results of this method of image analysis support the findings of the rigid body registration ROI analysis, which demonstrated that volume change over time is most

pronounced in sub-cortical structures in PSP and MSA-P. As suspected these changes do not simply occur at tissue boundaries. The regions affected correspond well with the distribution of the known histopathological changes seen at post mortem.

In the MSA-P cases, despite the predominant parkinsonian phenotype, marked tissue volume loss was demonstrated in the cerebellum confirming that the pathological process affects the cerebellum, supporting the results outlined using different techniques in chapters 6 and 9. The middle cerebellar peduncle was one of the most dramatically affected areas, supporting the findings reported in Chapter 8.

Visual assessment of the fluid registered images gives some information as to where regional atrophy is greatest, without any a priori assumptions regarding the regions concerned. In order to compare disease groups and identify changes that are significantly different, groups of patients need to be compared using robust statistical techniques. SPM can be applied to fluid registered MRI scans in order to achieve this.

11. An unbiased analysis of regional atrophy rates

11.1. Introduction

So far, the analysis of the longitudinal imaging findings has been based around hypothesis driven region of interest analysis. Based on knowledge from previous imaging and pathological studies, specific brainstem and lobar regional volumes have been measured and atrophy rates calculated in these regions.

An alternative approach is to use automated techniques that can accurately localise regional atrophy in an unbiased manner. If repeat scans are matched to baseline scans using voxel –level deformation fields based on a viscous fluid model as in the previous chapter, then the Jacobian determinants from the model can be used to quantify change at the voxel level and colour overlays can be used to illustrate this qualitatively on an individual level. To determine changes that differentiate one disease from another and from healthy controls, groups of individuals can be compared using SPM (Friston, 1995). This combination of SPM and fluid registration analysis has been used successfully in Alzheimer's disease to demonstrate significantly increased rates of hippocampal atrophy in pre-symptomatic and mildly affected patients (Scahill *et al.*, 2002).

An understanding of the regional pattern and progression of atrophy in PSP and MSA-P may be useful and an un-biased fluid-SPM analysis may highlight new regions of increased atrophy. If applied to monitor atrophy rates in clinical trials, it may be able to highlight differing patterns of atrophy occurring in treated and placebo groups of patients.

Combined fluid-SPM analysis has therefore been applied to determine significant group differences in atrophy between PSP, MSA-P, PD and healthy controls, without

making assumptions regarding the sites of atrophy. The fluid-SPM technique has also been applied to investigate whether atrophy patterns are different in groups of PSP patients with differing clinical features.

11.2. Methods

11.2.1. MRI acquisition

T1 weighted MRI scans were acquired at two time points as previously described. Inhomogeneity artefact was corrected for using DBC (Lewis and Fox, 2004) and as previously described, the brain was outlined using a semi-automated technique following which a 9 *dof* registration was performed to align the repeat scan onto the baseline image (Freeborough *et al.,* 1996). A non-linear fluid registration was then carried out (chapter 10). The Jacobian value produced by this model provides a measure of the volume change occurring at each voxel within the image.

Voxel compression maps were visually inspected to ensure adequate image quality before being entered into the group comparison as some images may have been marred by artefact.

11.2.2. Statistical MRI analysis

Analysis was performed using SPM-99 (Wellcome department of cognitive neurology, London) running in MATLAB (Mathworks incorporated, Sheraton, Massachusetts). The Jacobian images were spatially normalised to a customised template in standard stereotactic space. Contraction and expansion images were separated based on Jacobian values alone producing voxel expansion (value >1.0) and contraction (value <1.0) images. This was performed, as without it smoothing would remove any change detected (as expansion and contraction occur close to each other) and would effectively cancel each other out. Images were then smoothed with an isotropic Gaussian kernel of 8mm full width half maximum.

A condition and covariate model was used for analysis with time interval as a covariate and age and sex as nuisance variables. The contraction and expansion images were masked with a generous brain mask in order to remove non-brain matter and define the analysis volume.

Significance levels were set at $p<0.001$ uncorrected and in order to determine the areas undergoing the greatest rates of change, a second analysis, with a significance level set at $p<0.05$, corrected for false discovery rate (FDR) multiple comparisons was performed.

The PSP patients were divided into approximately equal groups on the basis of the median (50^{th} centile) FAB score, MMSE score, gaze palsy score and disease duration (years). The two groups were then compared independently with healthy controls and with each other, in order to determine whether atrophy in specific areas was associated with poorer performance on the FAB or MMSE, a more severe gaze palsy or longer disease duration.

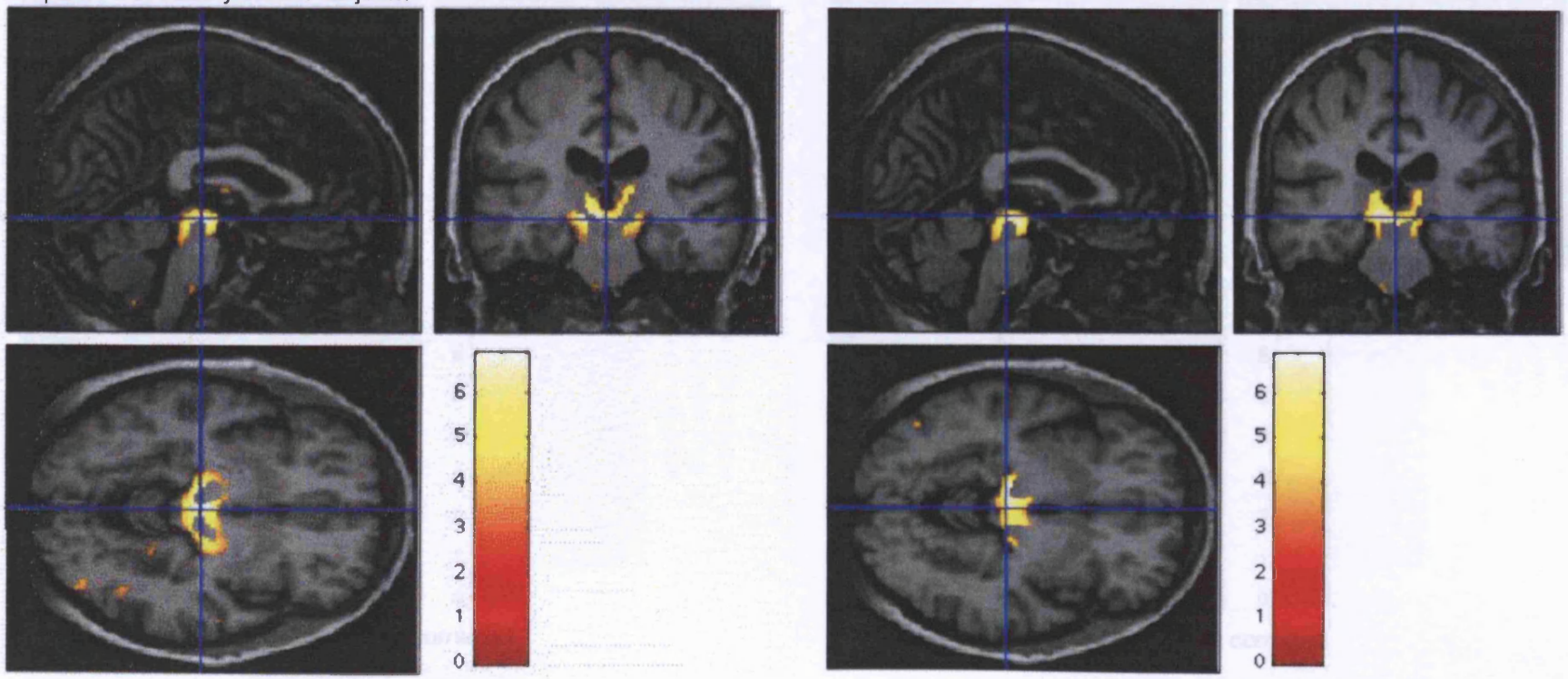
11.3. Results

Scans from 16 patients with PSP, 9 with MSA-P, 9 with PD and 12 healthy controls were included in the analysis after visual inspection for artefact on the fluid registrations was made. The clinical characteristics of the patients are described in table 11.1.

Table 11.1 *Mean (SD) characteristics of patients in fluid-SPM analysis.*					
	PSP	**MSA-P**	**PD**	**HC**	**P**
n	16	9	9	12	
Age	65 (5.9)	62.4 (8.1)	64.4 (8.3)	64.8 (5.3)	0.8
Disease duration	4.3 (1.4)	5.4 (1.7)	12.9 (4.3)	-	PSP/MSA-P vs. PD, p<0.001
Scan interval	296 (58.9)	262 (65.5)	268 (51.5)	254 (39.4)	PSP vs. HC, p=0.04
MMSE	26.4 (2.7)	25.9 (3.4)	28.2 (2.5)	29.5 (0.7)	PSP vs. HC, p<0.001, vs. PD, p=0.04, MSA-P vs. HC, p=0.004.
FAB	12 (2.9)	14.6 (2.8)	17 (1.1)	-	PSP vs. PD p<0.001, vs. MSA-P, p=0.05, MSA-P vs. PD, p=0.03
HY	3.4 (0.7)	3.8 (0.8)	2.7 (0.5)	-	p=0.006
UPDRS II	18.6 (6.8)	24.1 (7.6)	13.4 (4.6)	-	p=0.006
UPDRS III	19.1 (7.8)	24.7 (9.5)	16.4 (3.1)	-	NS
Gaze Palsy	12 (5.8)	2.5 (2.8)	-	-	p<0.001
Post hoc tests: HY- PSP vs. PD, p=0.04, MSA-P vs. PD, p=0.006; UPDRS II- MSA-P vs. PD, p=0.005.					

Figures 11.1 and 11.2 show statistical parametric maps in PSP and MSA-P, compared to healthy controls demonstrating regions of significantly increased rates of atrophy (tissue contraction). In the PSP patients, atrophy rates are significantly increased in the midbrain, cerebral peduncles and in the region of the medial thalamic nuclei. These regions survive a stricter statistical test of significance (FDR corrected $p<0.05$). In the MSA-P patients, significantly increased atrophy rates are seen in the pons, the midbrain, the cerebellar white matter and middle cerebellar peduncles. The thalamus in contrast to PSP is spared. These regions survive the stricter statistical test and figure 11.2 B also shows that greater rates of contraction compared to healthy controls are also seen in the medial and superior frontal gyri and in the temporal lobes. Increased rates of atrophy in MSA-P, compared with healthy controls are also clearly demonstrated in the cingulate gyrus, the orbitofrontal gyri, the gyrus rectus and the inferior rostral gyrus in figure 11.3.

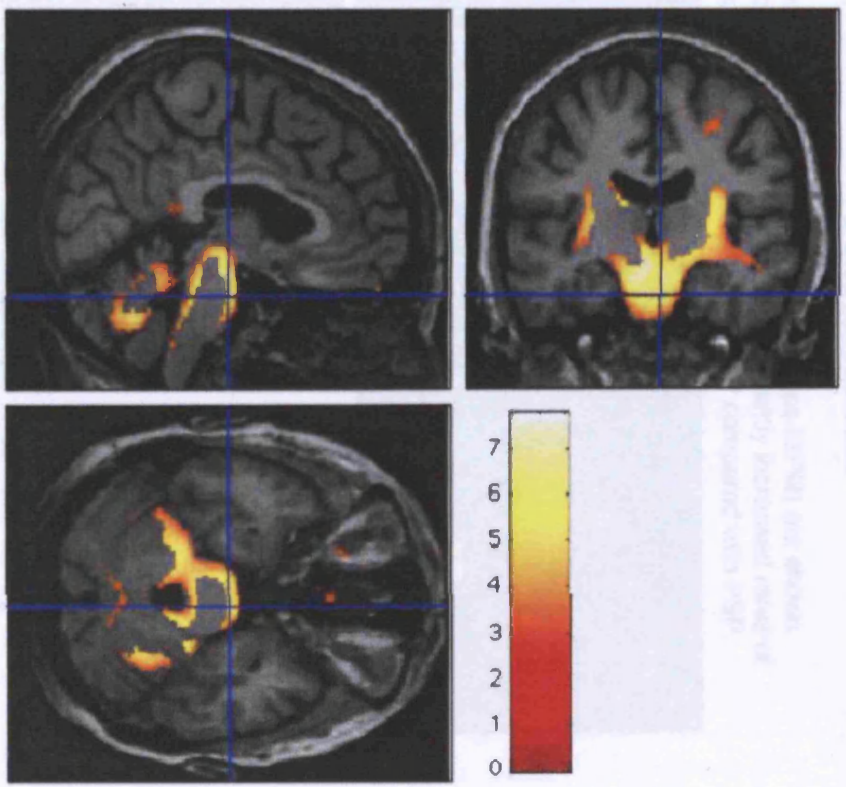
Figure 11.1 Patterns of regional atrophy in PSP.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of atrophy (tissue contraction) in PSP compared with healthy control subjects.



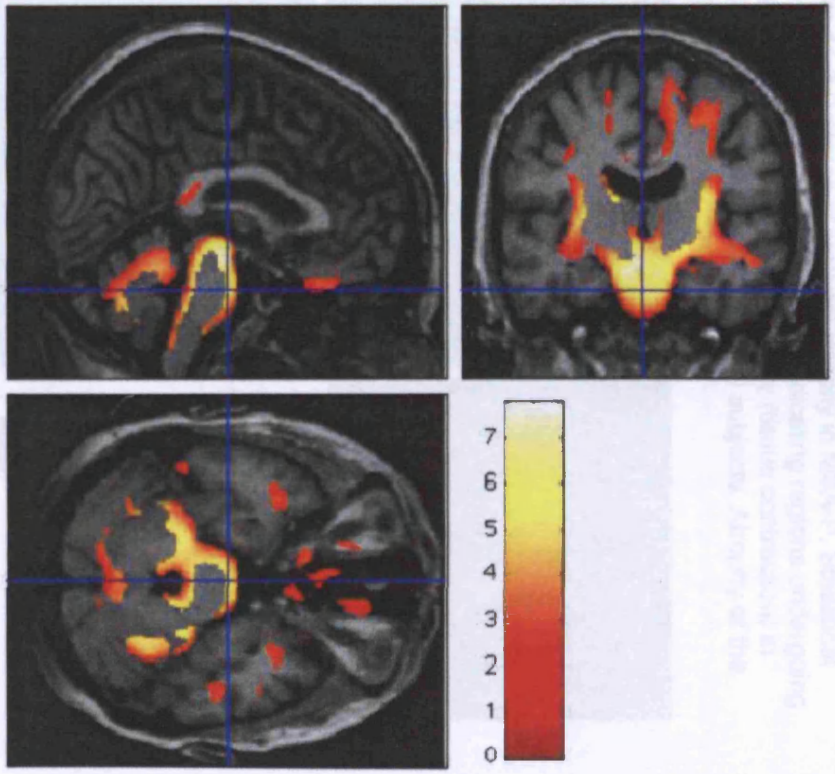
A p<0.001 uncorrected

B p<0.05 FDR corrected

Figure 11.2 Patterns of regional atrophy in MSA-P.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of atrophy (tissue contraction) in MSA-P compared with healthy control subjects.

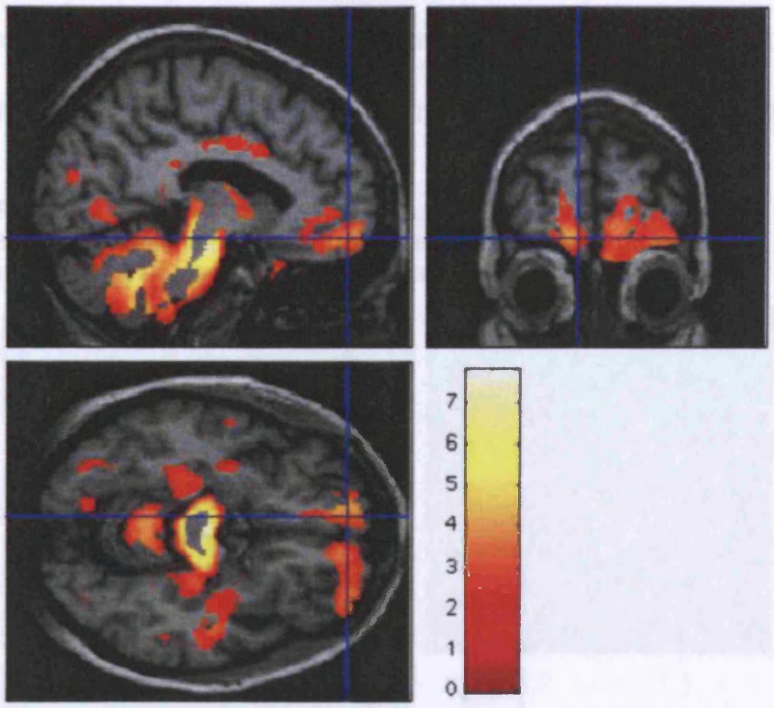


A p<0.001 uncorrected



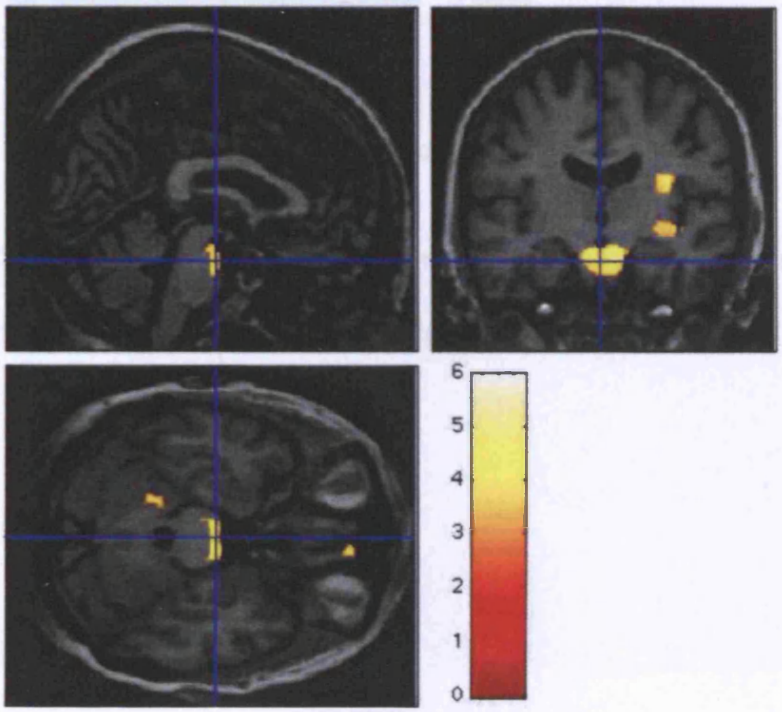
B p<0.05 FDR corrected

Figure 11.3 Patterns of regional atrophy in MSA-P. Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of atrophy (tissue contraction) in MSA-P compared with healthy control subjects. Atrophy of the orbitofrontal cortex is demonstrated.



p<0.05 FDR corrected

Figure 11.4 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of atrophy (tissue contraction) in MSA-P compared with PSP.



p<0.05 FDR corrected

Figure 11.4 shows regions undergoing significantly increased rates of tissue contraction in MSA-P compared with PSP. A significantly increased rate of atrophy in the pons and cerebellum as well as the temporal stem and insular cortex in MSA-P is the only clear significant difference. SPM analysis of atrophy patterns in PSP compared to MSA-P did not demonstrate any clear differences after FDR correction.

Figure 11.5 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of atrophy (tissue contraction) in PD compared with healthy controls.

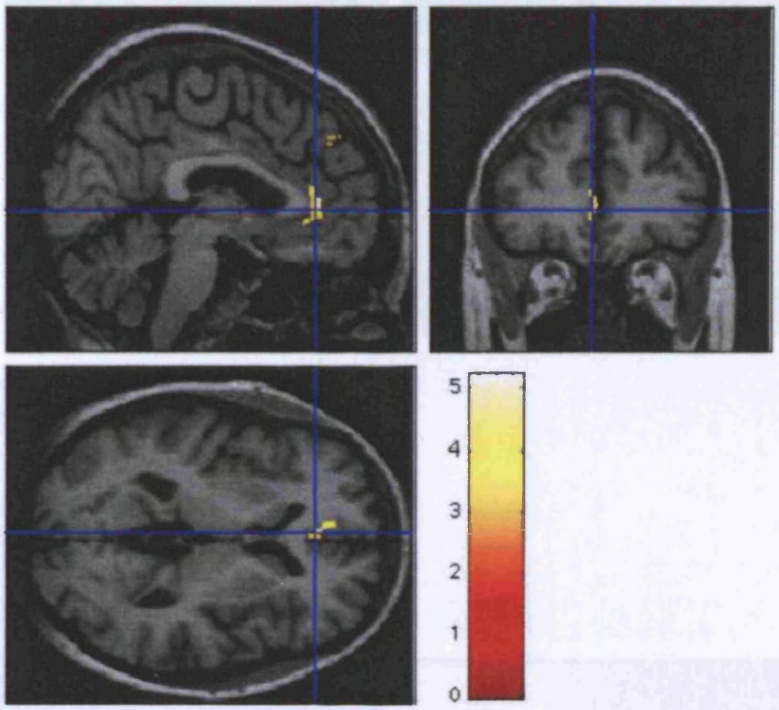
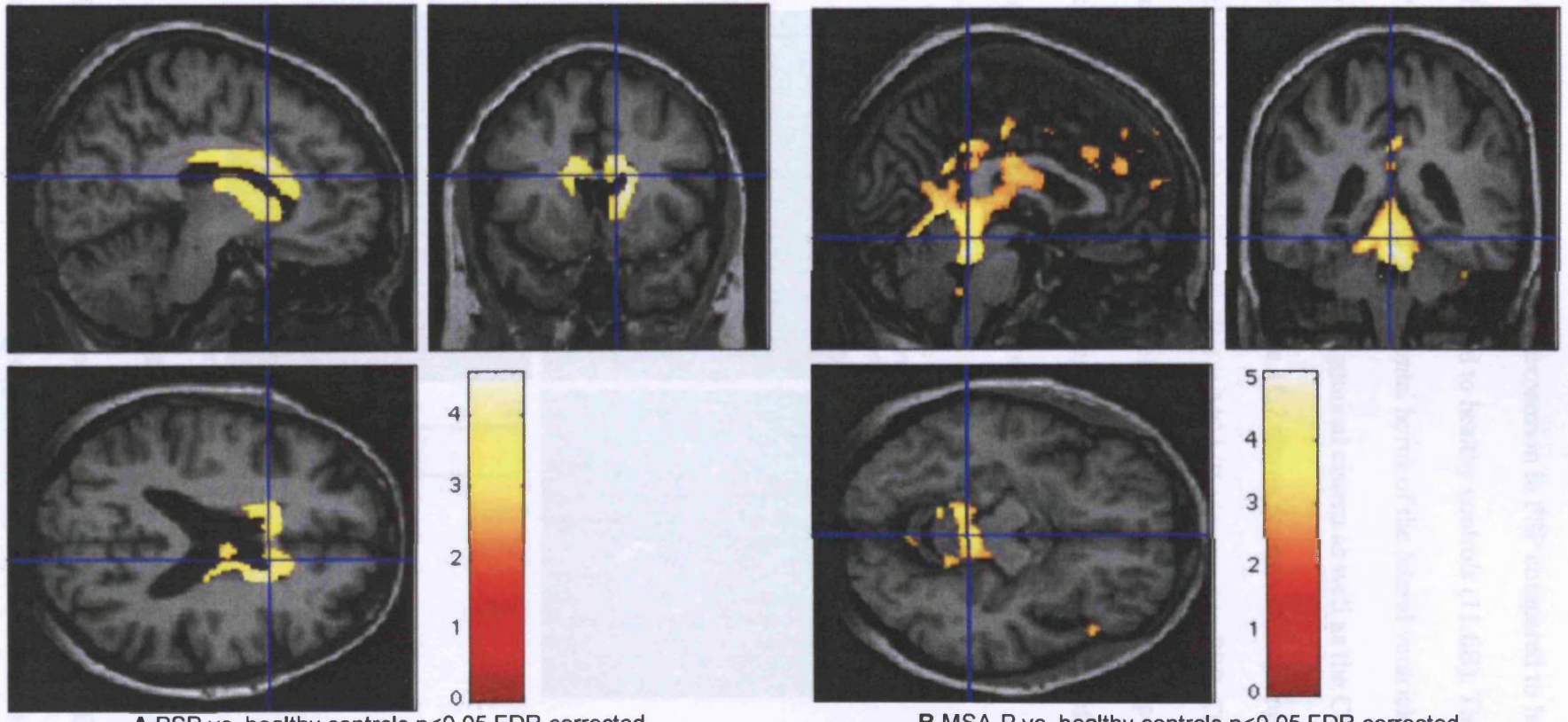


Figure 11.5 shows that in PD compared to healthy controls, only a minimal increase in atrophy rates in the region of the cingulate gyrus. This region did not survive FDR correction ($p<0.05$)..

Figure 11.6 Patterns of regional tissue expansion in PSP and MSA-P.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in PSP (A) and MSA-P (B) compared with healthy control subjects.

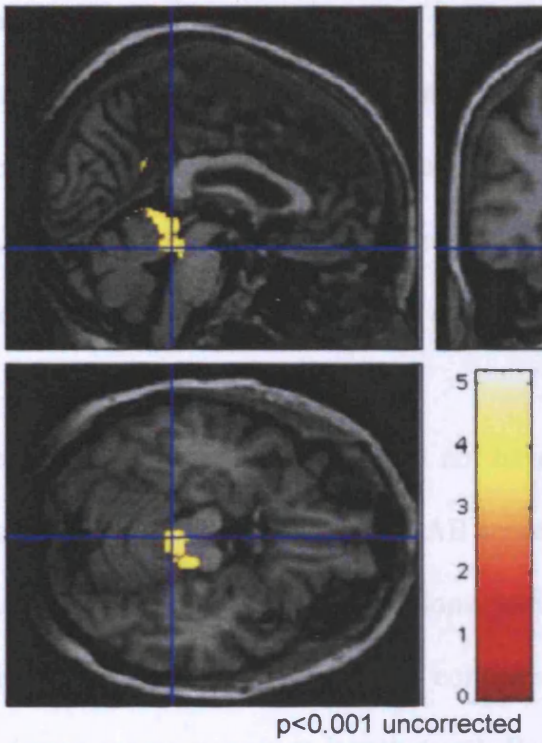


A PSP vs. healthy controls p<0.05 FDR corrected

B MSA-P vs. healthy controls p<0.05 FDR corrected

Figure 11.6 shows regions of significant expansion in PSP compared to healthy controls (11.6A) and in MSA-P compared to healthy controls (11.6B). The differences in the PSP patients are confined to the frontal horns of the lateral ventricles and in MSA-P to the fourth ventricle and quadrigeminal cistern as well as the CSF space around the cerebellar hemispheres. Figure 11.7 shows regions undergoing significantly increased rates of expansion in MSA-P compared to PSP. The fourth ventricle and quadrigeminal cistern as well as CSF spaces around the superior aspect of the cerebellum are the most affected regions. There were no regions of significantly increased rates of expansion in PSP compared to MSA-P

Figure 11.7 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in MSA-P compared with PSP.



p<0.001 uncorrected

11.3.1. The influence of FAB, MMSE, gaze palsy and disease duration

Table 11.2 shows the division of PSP subjects based on FAB, MMSE, gaze palsy scores and disease duration. The subgroups of PSP patients were compared with

healthy controls and with each other using age and sex as nuisance variables and scan interval as a covariate in the SPM analysis.

Table 11.2 *Subgroups based on median scores in PSP*

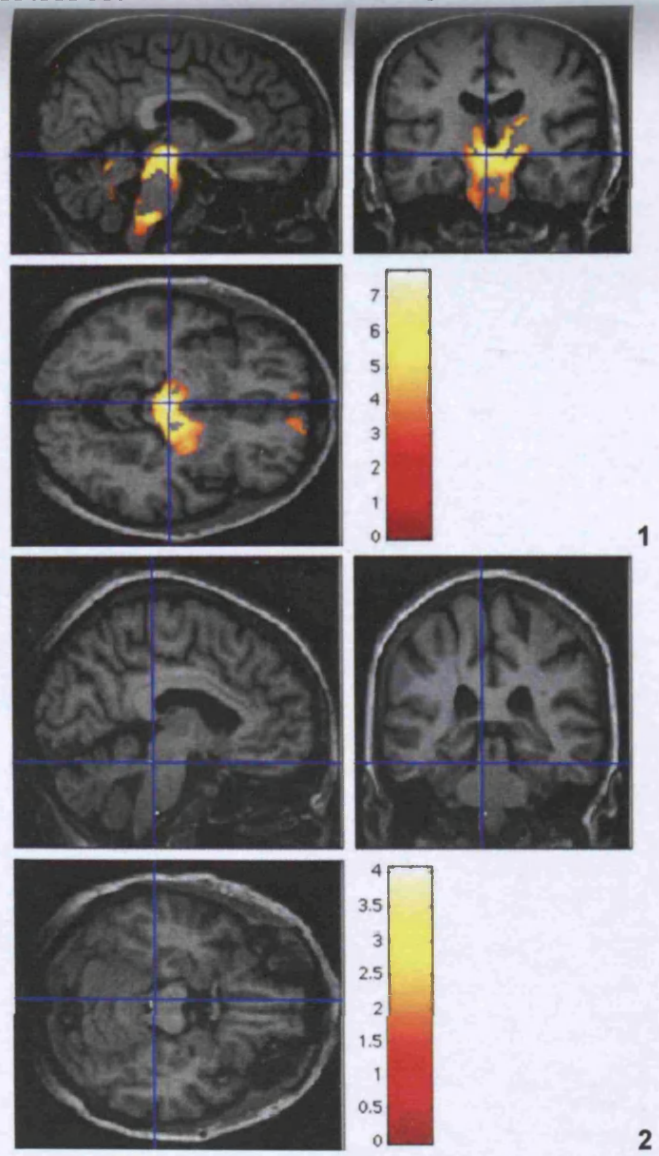
	Median score		PSP n
FAB	12	≥ 12	7
		<12	9
MMSE	27	≥ 28	7
		<28	9
Gaze palsy	12	≥ 12	9
		<12	7
Disease duration	4.5	>4.5	8
		<4.5	8

FAB: Dividing the PSP patients according to the FAB separated the group into two subgroups of n=7 and n=9. There was no difference in mean disease duration (p=0.14) or age (p=0.87) between them. SPMs showing regions undergoing significantly increased rates of contraction and expansion are shown in figures 11.8 and 11.9.

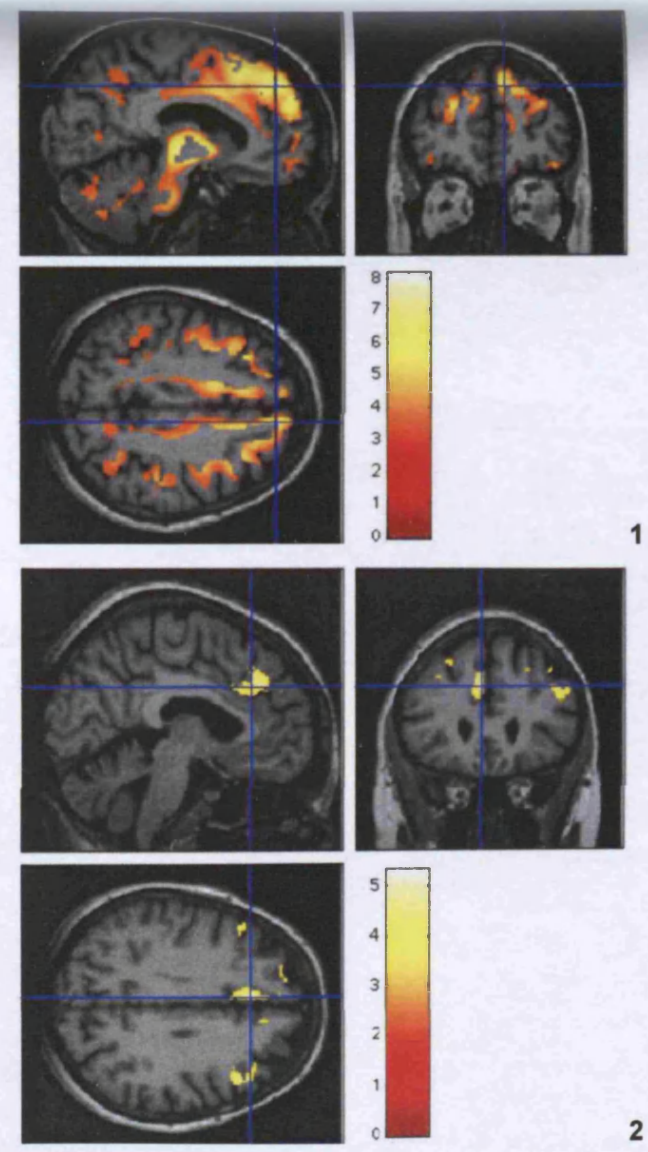
Figure 11.8 (B1) demonstrates that patients with a high FAB score (less frontal/subcortical cognitive impairment) have significantly increased rates of tissue contraction, bilaterally in the superior and middle frontal gyri as well as the orbitofrontal gyrus, the cingulate cortex, the pre-central gyrus and the superior parietal regions.

Those patients with a lower FAB score did not have any areas of increased atrophy rates compared to the group with a higher FAB score. Those with a higher FAB score had a greater rate of atrophy in frontal regions, particularly the superior and middle frontal gyri and the cingulate gyrus, when compared directly to those with a lower FAB score (more severe cognitive impairment). Rates of expansion were greater in those with a higher FAB score compared to those with a low FAB score (figure 11.9).

patients with a low and a high FAB score compared with healthy controls (A1 and B1) and each other (A2 and B2). A higher FAB score indicates *less severe* frontal/sub-cortical cognitive impairment.

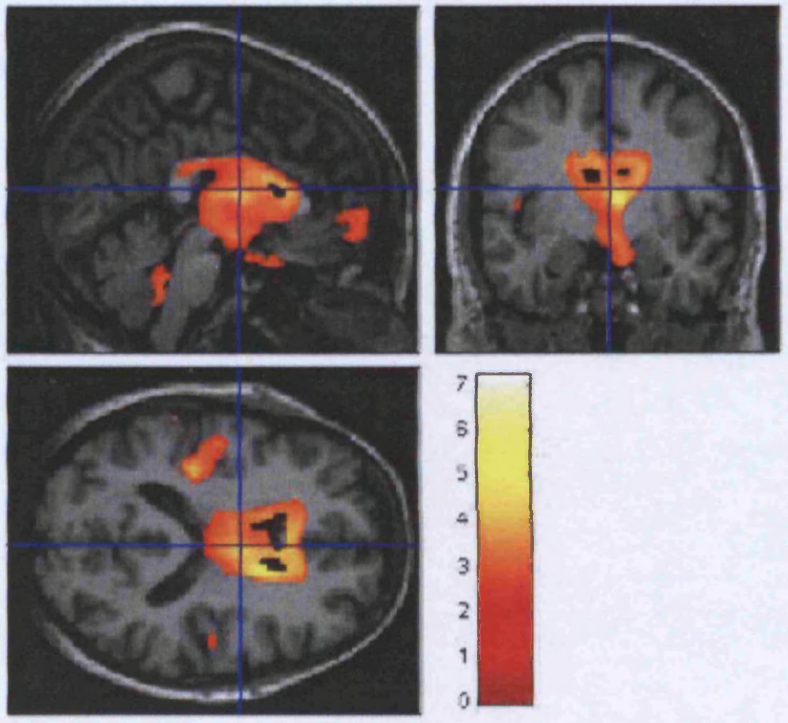


A-PSP with low FAB score(<12) compared to healthy controls (1) and to PSP patients with a high FAB (≥12) (2) -p<0.001 uncorrected



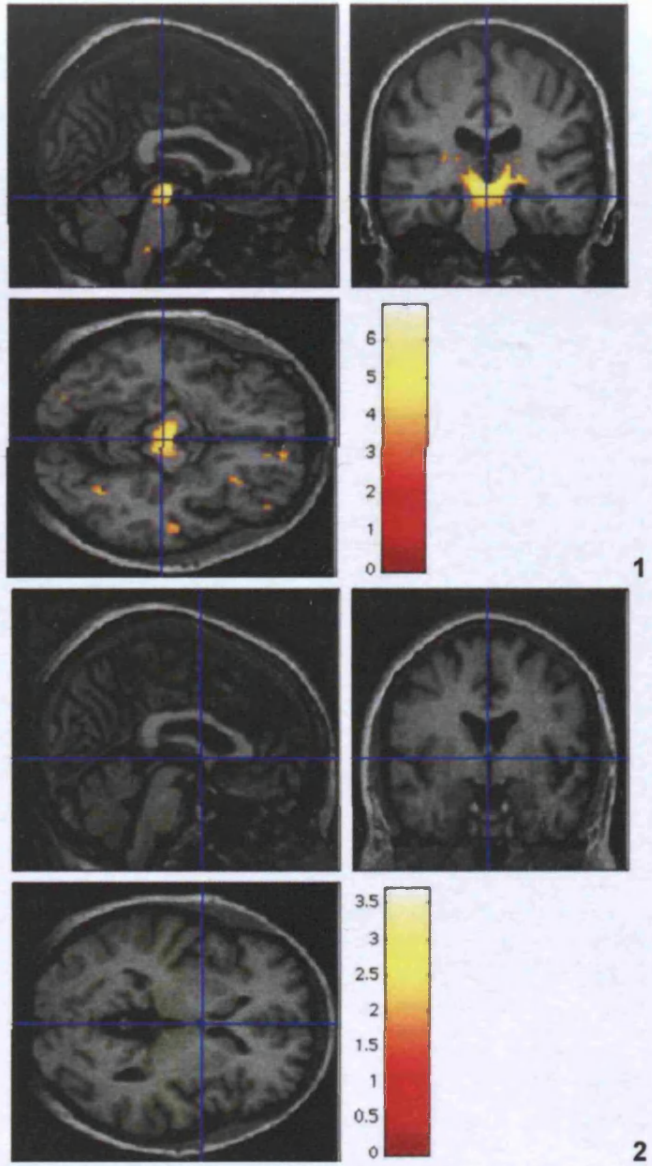
B-PSP with high FAB score (≥12) compared to healthy controls (1) and to PSP patients with a low FAB (<12) (2) -p<0.001 uncorrected

Figure 11.9 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in PSP (FAB ≥12) vs. PSP (FAB<12).

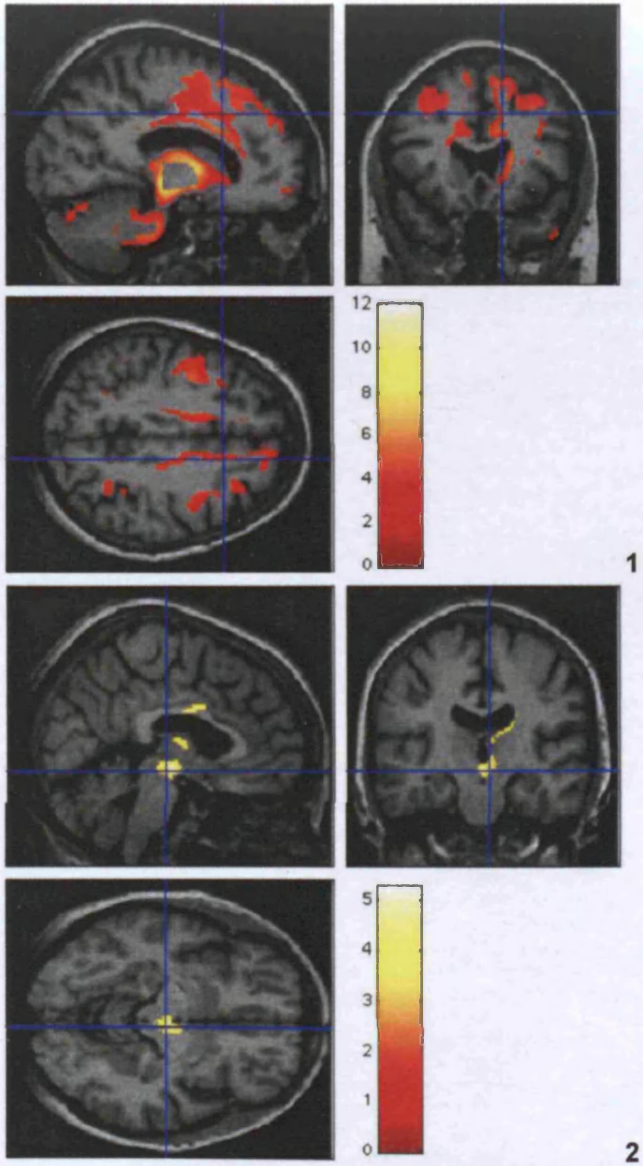


p<0.05 FDR corrected

Figure 11.10 Patterns of regional tissue contraction.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of contraction in PSP patients with a disease duration of >4.5 and <4.5 compared with healthy controls (A1 and B1) and each other (A2 and B2).

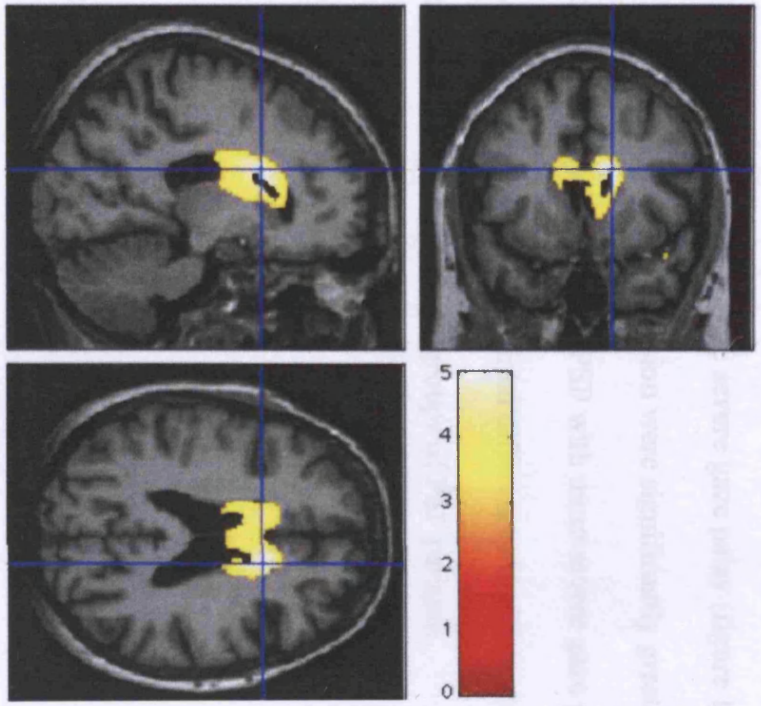


A-PSP with disease duration >4.5 compared to healthy controls (1) and to PSP patients with disease duration <4.5 (2) -p<0.001 uncorrected



B- PSP with disease duration <4.5 compared to healthy controls (1) and to PSP patients with disease duration >4.5 (2) -p<0.001 uncorrected

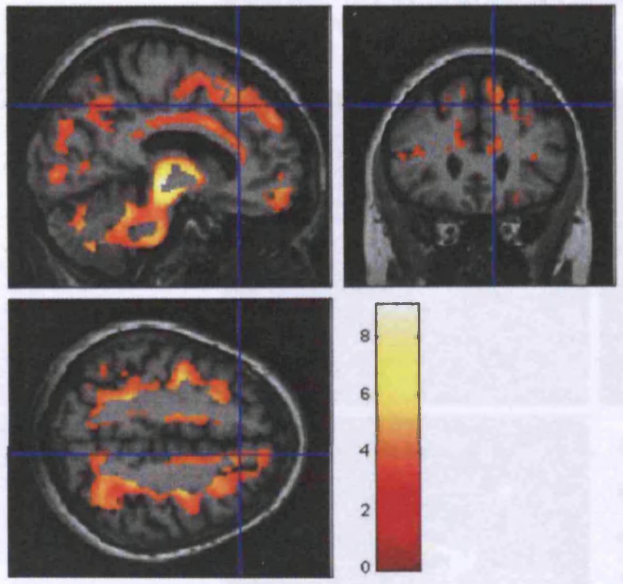
Figure 11.11 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in PSP (disease duration <4.5 years) vs. PSP (disease duration >4.5 years).



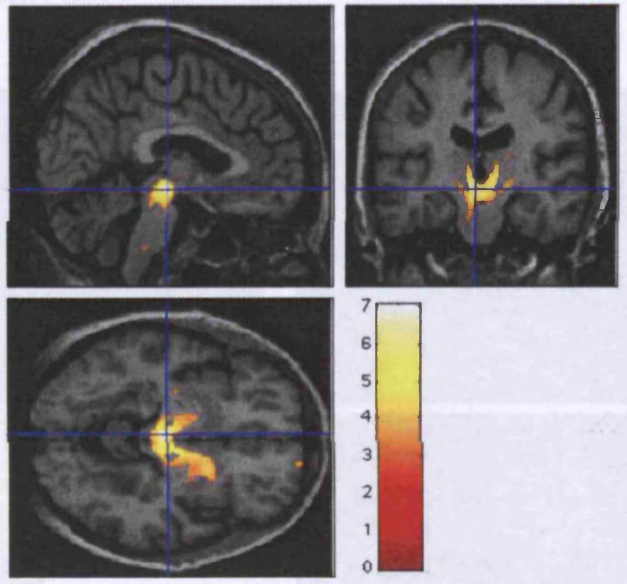
p<0.05 FDR corrected

Gaze palsy: Dividing the PSP patients into two groups according to gaze palsy score produced two groups (n=7 and n=9) with no difference in mean disease duration (p=0.78) or age (p=0.48). SPMs show greater rates of atrophy in those patients with a more severe gaze palsy in the superior, medial and orbitofrontal gyri as well as the parietal cortex and brainstem regions, when compared to healthy controls (figure 11.12 A). In PSP patients with a less severe gaze palsy, the significantly different regional rates of atrophy were confined to the brainstem (figure 11.12 B). Comparing the two sub-groups of PSP patients, showed significantly increased rates of atrophy in the superior parietal lobule in those with a more severe gaze palsy (figure 11.12 C). Figure 11.13 A and B show that rates of expansion were significantly greatest in the lateral ventricles and over the frontal cortex in PSP with more severe gaze palsy compared to healthy controls but that only minimal differences in regional expansion persisted when directly comparing the two subgroups of PSP patients.

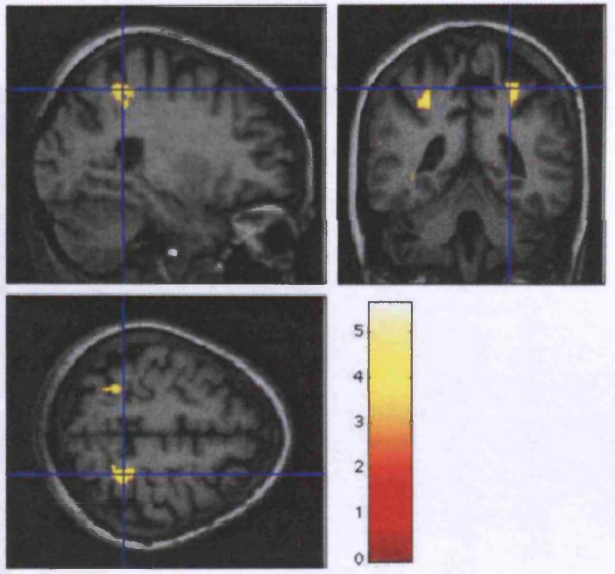
Figure 11.12 Patterns of regional tissue contraction.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of contraction in PSP patients with a low and a high gaze palsy score compared with healthy controls and each other. A higher gaze palsy score indicates a *more severe* gaze palsy.



A- PSP (gaze palsy≥12) compared to healthy controls-p<0.001 uncorrected

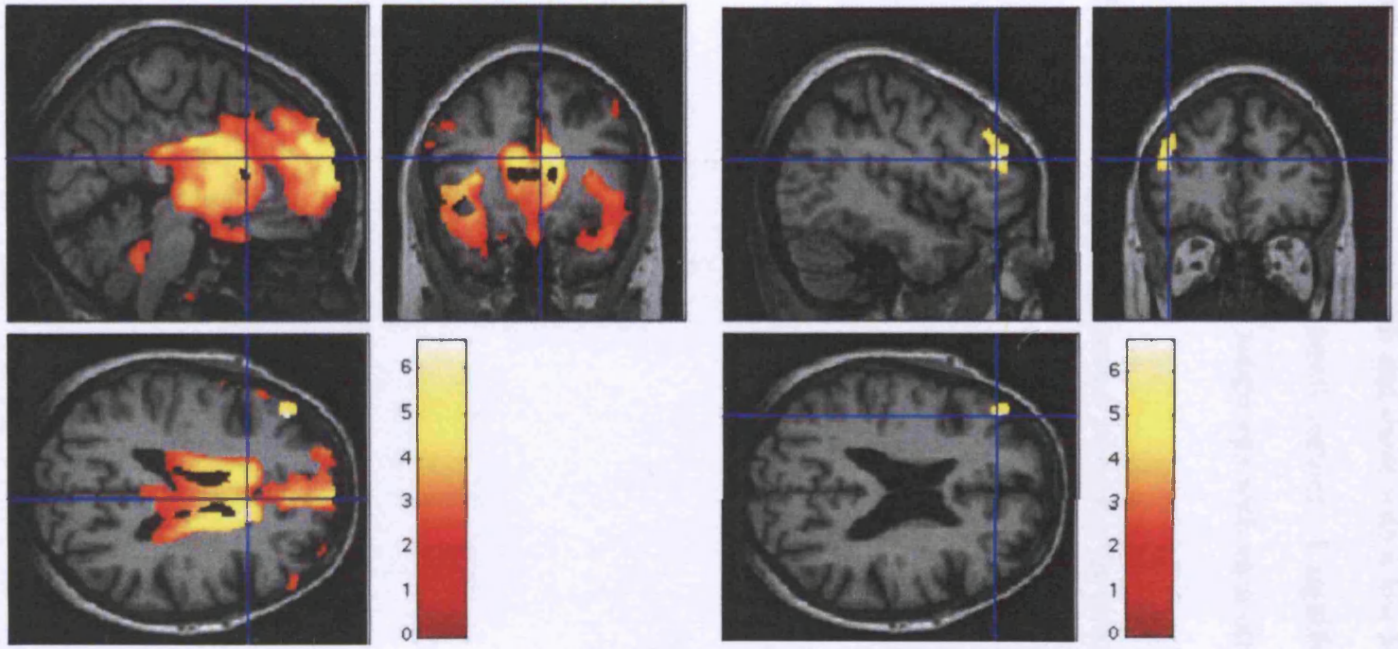


B- PSP (gaze palsy<12) compared to healthy controls-p<0.001 uncorrected



C- PSP (gaze palsy≥12) compared to PSP (gaze palsy<12) p<0.001 uncorrected

Figure 11.13 Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in PSP with severe gaze palsy (score >12) compared to healthy controls (A) and to PSP patients with a gaze palsy score <12 (B).



A PSP (gaze palsy ≥12) vs. healthy controls p<0.05 FDR corrected

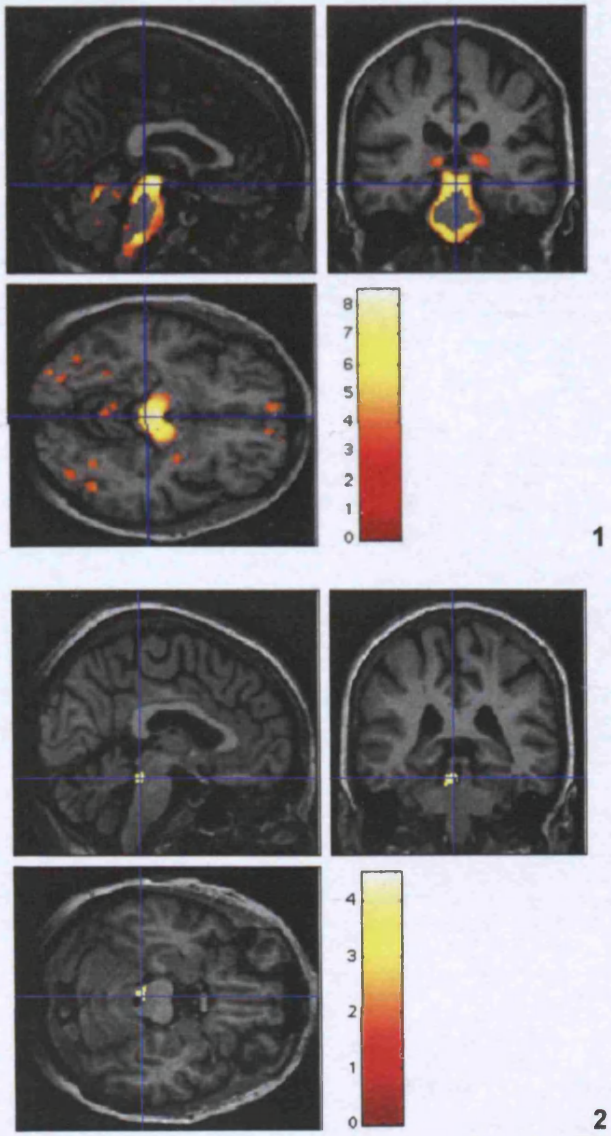
B PSP (gaze palsy ≥12) vs. PSP (gaze palsy <12) p<0.05 FDR corrected

MMSE: Dividing the PSP patients according to MMSE resulted in two groups (n=7 and n=9). There was no significant difference in age (p=0.59) or disease duration (p=0.18) between the two groups. The SPMs reveal similar patterns of significantly increased atrophy in those with a high MMSE and those with a low MMSE compared to healthy control subjects and only very small regions of significantly increased atrophy rates when comparing the two PSP subgroups with each other (figure 11.14 A1-B2).

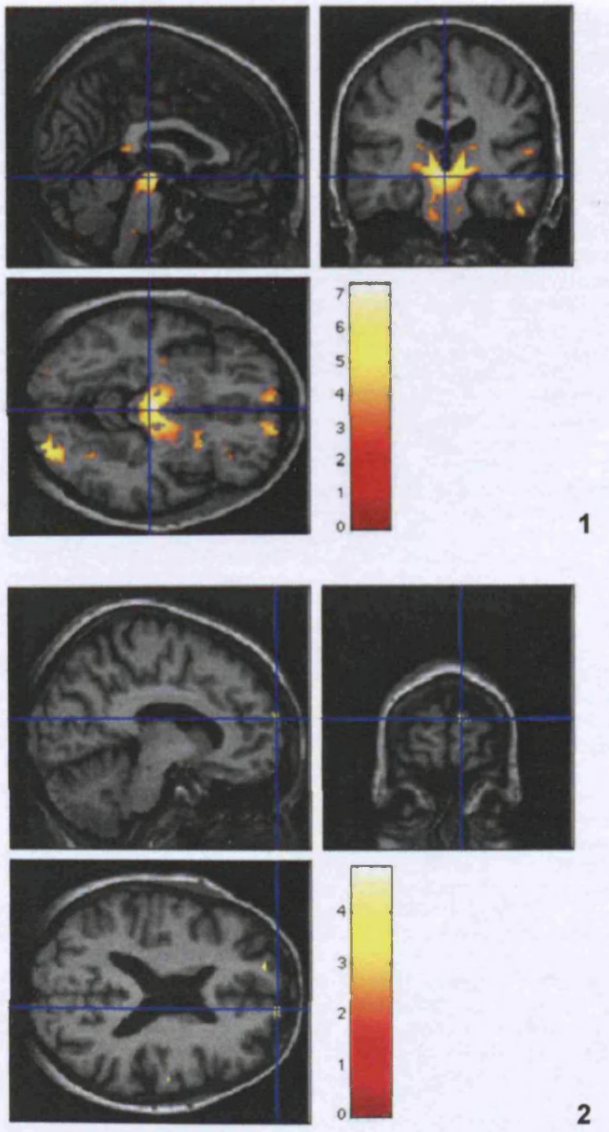
A similar pattern of significant regional expansion affecting the anterior horns of the lateral ventricles is also seen in both PSP subgroups compared to healthy controls (figure 11.15 A and B).

Figure 11.14 Patterns of regional tissue contraction.

Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of contraction in PSP patients with MMSE of <28 and, ≥ 28 compared with healthy controls (A1 and B1) and each other (A2 and B2).

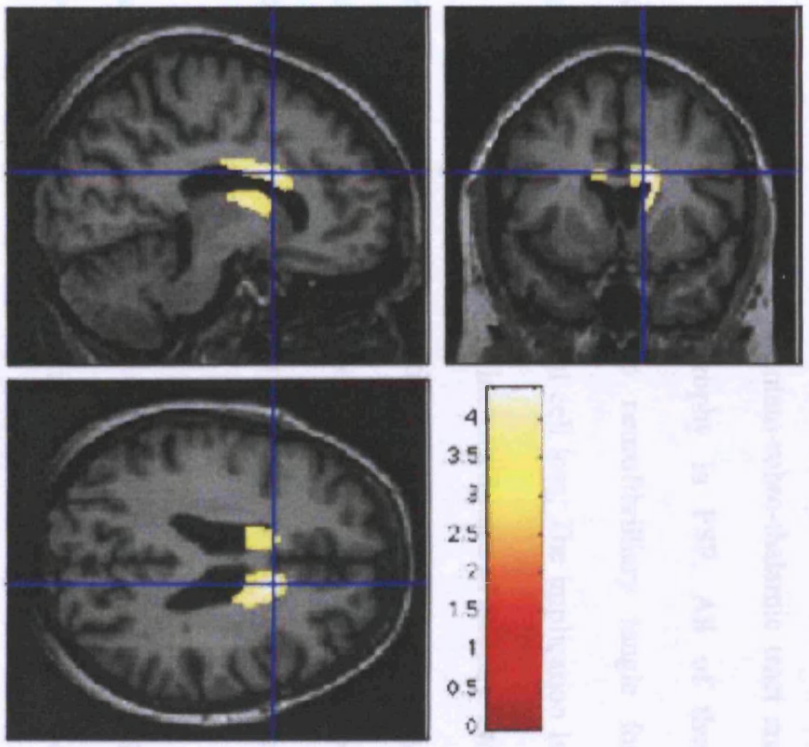


A-PSP with MMSE<28 compared to healthy controls (1) and to PSP patients with MMSE ≥ 28 (2) -p<0.001 uncorrected

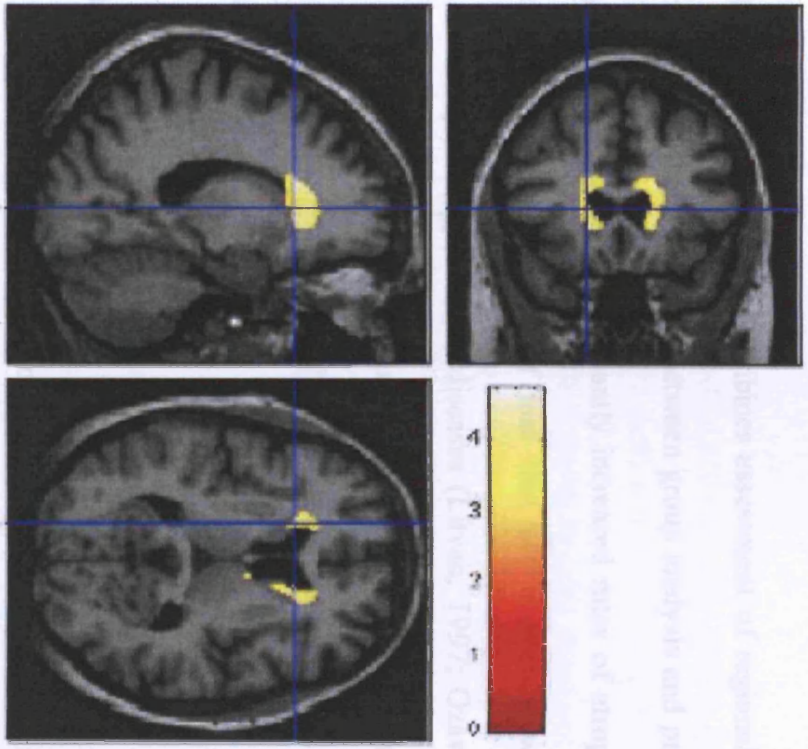


B- PSP with MMSE≥ 28 compared to healthy controls (1) and to PSP patients with MMSE <28 (2) -p<0.001 uncorrected

Figure 11.15 Patterns of regional tissue expansion in PSP.
Statistical parametric maps (SPM) are shown indicating regions undergoing significantly increased *rates* of expansion in PSP with MMSE <28 vs. healthy controls (A) and PSP with MMSE ≥ 28 vs. healthy controls (B).



A PSP (MMSE <28) vs. healthy controls p<0.05 FDR corrected



B PSP (MMSE ≥ 28) vs. healthy controls p<0.05 FDR corrected

11.4. Conclusions

The technique used in this chapter combines assessment of regional atrophy on an individual basis with robust statistical between group analysis and provides a means of assessing areas that undergo significantly increased rates of atrophy in PSP and MSA-P. The results indicate areas of loss that are consistent with the known distribution of histopathology in these diseases (Litvan, 1997; Ozawa *et al.*, 2004; Papp and Lantos, 1994). Importantly, this data is in keeping with the results of the region of interest analysis described in chapters 6 and 9, and the cross sectional VBM analysis described in chapter 7.

PSP compared to healthy controls

Significantly increased rates of tissue contraction in PSP compared to healthy controls were present in the midbrain and cerebral peduncles. The substantia nigra, subthalamic nucleus, red nuclei and dentato-rubro-thalamic tract are present in the regions of significantly increased atrophy in PSP. All of these regions are significantly affected by tau positive neurofibrilliary tangle formation, tufted astrocytes, neuropil threads and neuronal cell loss. The implication is then that these histological changes are associated with increased rates of atrophy. Pathology in these regions is implicated in the expression of the typical clinical phenotype of PSP with early falls and an akinetic rigid syndrome. The typical gaze palsy associated with PSP also probably arises as a consequence of pathological involvement of the upper brainstem.

No significant increase in frontal atrophy rates was detected in the PSP patients, when the whole PSP group was compared to the controls. This may be because the whole

group includes some subjects with increased rates of loss and a number without increased rates of frontal atrophy.

Cross sectional VBM analysis demonstrated reduced grey matter voxel intensity in the medial frontal gyrus and the insular cortex in the PSP patients compared to healthy controls. It is possible that no increase in atrophy rates was detected in the PSP group as a whole in this region because the rate of tissue contraction had already reached a plateau in the more advanced cases. A previous cross sectional VBM study helps to support this theory as it demonstrated a reduction in frontal volume in a group of PSP patients with a mean disease duration of only 2.7 years, compared to that of 4.3 years in those recruited to this study. However the potential for ongoing frontal atrophy is large and that atrophy rates simply arrest after several years of disease seems unlikely.

It is more probable that the PSP group contains patients who have increased rates of frontal atrophy as well as those who do not, and overall, no significant difference is detected.

Increased rates of atrophy in the thalamus were detected in PSP compared to healthy controls (figure 11.1B and 11.2B) supporting the findings of the DWI analysis. A recent study of thalamic pathology in PSP confirmed that the medial thalamus contains large numbers of NFTs (Henderson *et al.*, 2000). Whereas, the limbic thalamus is less affected. Atrophy in the centro-median nucleus of the thalamus in PSP has been suggested as reflective of thalamic deafferentation due to degeneration of basal ganglia output in PSP (Hardman *et al.*, 1997). Whatever the mechanism behind atrophy of the median thalamus, it is likely that disruption to thalamic relays, either as a consequence of direct pathological involvement or deafferentation

secondary to extensive basal ganglia pathology contributes to the basal ganglia signs characteristic of PSP.

PSP compared to MSA-P

The fluid-SPM analysis did not detect any significant differences in atrophy rates between PSP and MSA-P subjects. This is in contrast to the results of the serial volumetric ROI analysis reported in chapter 9. It is possible that no group differences survived the strict statistical tests used in the fluid-SPM analysis; however the ROI analysis demonstrated that in addition to the midbrain atrophy rates being greater than controls in PSP as expected, they were also increased in MSA-P. This might explain why no significant increase in atrophy rates in this region was observed when comparing PSP and MSA-P. Atrophy in basal ganglia regions and even cortical regions may be expected in MSA-P explaining why increased rates of atrophy in these regions when comparing the two diseases were not detected.

MSA-P compared to healthy controls

Regional differences in atrophy rates between MSA-P and healthy controls were present in the pons and the midbrain, the cerebellum and the middle cerebellar peduncle. The thalamus appeared relatively spared. There were, however, regions of increased rates of cortical atrophy in MSA-P including the superior frontal gyri, the temporal lobes and the anterior cingulate cortex and orbitofrontal gyri. That increased rates of cortical atrophy occur in MSA-P is in keeping with post mortem studies demonstrating some frontal atrophy (Konagaya *et al.*, 1999b; Wakabayashi *et al.*, 1998). Although, these case reports comment on post mortem pathological involvement of motor cortical areas after disease durations of 13 or 14 years. Conventionally, frontal cortical atrophy detected on MRI during life would not have

been considered indicative of MSA-P and might even have been considered supportive of a diagnosis of PSP or CBD (Schrag *et al.*, 2000). The mean disease duration of MSA-P patients recruited for fluid-SPM analysis was 5.4 years and our results suggest that frontal atrophy is already ongoing at this stage in the disease process. This together with the results of a previous cross sectional VBM study in MSA-P (Brenneis *et al.*, 2003) and the results of our own cross sectional VBM study (chapter 7) confirm that atrophy affecting the insular cortex, the medial frontal gyrus, the anterior cingulate cortex and the dorsal pre-motor cortex occurs in MSA-P. These cortical regions are similar to those involved in MSA-C (Brenneis *et al.*, 2005). The phenotypic differences in MSA subtypes are a matter of relative rather than absolute differences in the severity of the pathological involvement (Ozawa *et al.*, 2004). As such, greater rates of cerebellar atrophy would be expected in MSA-C.

Cortical pathology in MSA has been demonstrated in the sensorimotor and higher cortical areas such as the supplementary motor area and the anterior cingulate cortex (Papp and Lantos, 1994). Case reports of frontal cortical atrophy in MSA have also demonstrated loss of small and medium sized pyramidal cells in the pre-central gyrus, superior and middle frontal gyri, with glial cytoplasmic inclusions in the deeper cortical areas resulting in extensive myelin and axonal loss in the frontal white matter. This supports the notion that in MSA, the increased MRI atrophy rates occur as a consequence of direct pathological involvement rather than secondary deafferentation. An obvious clinical correlate of frontal cortical atrophy is pyramidal tract dysfunction in MSA-P. This was present in all, except one, MSA-P patient recruited to this study. The MSA-P patients also demonstrated variable neuropsychological evidence of frontal cognitive impairment.

MSA-P compared to PSP

Regions of significantly increased rates of tissue atrophy in MSA-P compared to PSP are restricted to atrophy in the pons and in the cerebellum although the size of the regions that were statistically significantly different after correction for multiple comparisons is small. Increased atrophy rates in the region of the insular cortex in MSA-P was also identified suggesting that some regional cortical atrophy may be occurring at a greater rate in MSA-P than in PSP. Alternatively, it is possible that atrophy rates in these regions in PSP have fallen off as tissue contraction has already reached its maximum.

A number of conclusions can be drawn from the PSP subgroup analysis.

Frontal cortical atrophy seems to occur early in the course of the disease. In those with less severe cognitive impairment on the FAB, rates of frontal atrophy were greater than in those with more severe cognitive impairment. One potential explanation for this is that atrophy in frontal regions had already occurred in those with more severe cognitive impairment, but is ongoing in those with less severe cognitive impairment. The lack of more severe frontal volume reduction in PSP according to the cross sectional analysis (Chapter 7) may be due to varying rates of atrophy according to the disease phenotype. In the PSP patients with shorter reported disease duration, atrophy rates were greater in the same frontal regions implying it occurs early in PSP, and does not accelerate in the later stages of the disease.

Atrophy rates and the FAB

When divided according to the median FAB score in the PSP group, those with less frontal/subcortical cognitive impairment according to the bedside rating scale (FAB≥12) had significantly greater rates of atrophy in the brainstem, the superior and

medial frontal gyri, the orbitofrontal gyri, the cingulate cortex and the pre-central gyrus when compared to healthy controls. Those with lower FAB scores (<12) had increased rates of brainstem atrophy compared to controls but did not have increased rates of frontal cortical atrophy. Disease duration and age were no different between the two groups.

When comparing the groups directly, increased rates of atrophy persisted in the superior and medial frontal gyri in those with a higher FAB score. This suggests that atrophy in these regions has already occurred in those with more severe cognitive impairment and in those with less severe cognitive impairment atrophy in frontal cortical regions is still ongoing. That these frontal regions are implicated is not surprising as the FAB test score is dependent on functions mediated by the pre-frontal cortex (Chapter 5.4)

The regions of significantly increased atrophy correspond with those identified in previous cross sectional imaging studies in PSP (Brenneis *et al.,* 2004). These regions are targets of striatal projections and are engaged in regulation of motor initiation, response selection, motivation and goal directed behaviour. Hence that increased atrophy rates in the superior and medial frontal cortex in PSP is linked with FAB score is in keeping with the apathy, impaired executive function and action initiation as well as difficulty with set shifting consistently demonstrated on neuropsychological testing in PSP.

These findings support the conclusion that atrophy as a consequence of direct pathological involvement of frontal cortical structures as well as secondary to deafferentation contributes to the frontal/subcortical dementia syndrome of PSP.

Atrophy rates and disease duration

When divided according to disease duration, increased rates of atrophy were detected in frontal cortical areas as well as brainstem regions in those with a shorter disease duration compared to healthy controls. Frontal cortical atrophy rates were not significantly increased when comparing those with longer disease duration to healthy controls. This suggests that cerebral atrophy in PSP occurs early in the course of the disease. When directly comparing PSP patients with a short and long disease duration, greater rates of atrophy were only identified in the brainstem regions in those with disease duration of less than 4.5 years. This still supports the theory that rates of atrophy decline as the disease progresses. This has important implications if MRI is to be considered as a means of monitoring disease progression in drug trials and indeed when considering a disease modifying therapy in the later stages of the disease. Firstly, in a drug trial, a real but small, drug effect may be missed if patients with longer disease duration are recruited as their atrophy rates are slower hence a greater number will be needed to detect a treatment effect (see chapter 14 for further discussion). Secondly if macroscopic atrophy occurs as a consequence of neuronal cell loss and this has slowed in the later stages of the disease, it implies that a disease modifying therapy may only have a meaningful effect early in the course of the disease when atrophy rates are greatest and there is more neuronal function to preserve.

That atrophy rates are greater early in the course of the illness suggests that application of MRI derived regional atrophy rates as a means of supporting a clinical diagnosis of a neurodegenerative disease is valid. This technique could also readily be applied to monitor disease progression early in the course of the illness.

These findings are in contrast to fluid-SPM studies in Alzheimer's disease, which conclude that atrophy rates accelerate later in the disease course (Scahill *et al.*, 2002). So far, these conclusions assume that all PSP patients have similar rates of disease progression and the greater atrophy rates detected in those with a disease duration of <4.5 years reflect that atrophy rates slow as the disease progresses. An alternative explanation is that those with a longer disease duration have survived longer (>4.5 years) as they have a slower rate of cell death.

Atrophy rates and gaze palsy

Comparing PSP patients divided according to the severity of gaze palsy suggests that greater rates of cerebral and brainstem atrophy occur in those patients with a more severe gaze palsy when compared with healthy controls. When compared directly with each other, increased rates of atrophy were identified bilaterally in the superior parietal lobe, in the region of the supra marginal gyrus suggesting that ongoing atrophy in this region occurs in those with a more severe gaze palsy. The parietal eye fields project to the superior colliculus and from there to the brainstem saccadic generators, and are important in the triggering of reflexive visually guided saccades. There is evidence to suggest that the parietal eye fields accumulate sensory signals relevant to the selection of a target for an eye movement and in humans, the latency of visually triggered saccades is increased when lesions occur in this region. Transcortical magnetic stimulation of the posterior parietal cortex supports the view that it contributes to reflexive saccades (Leigh and Kennard, 2004). Clearly in PSP, the pathological involvement of brainstem regions responsible in part for the generation of saccadic eye movements means that more than just difficulty with reflexive saccades occurs. Even in those with less severe gaze palsy, tissue atrophy is

greater than in healthy controls in the brainstem. However pathological involvement of the parietal eye fields has been demonstrated to be less severe in PSP cases with no documented gaze palsy during life, confirming a pathological basis for the VBM findings demonstrated in this thesis.

Greater cerebral atrophy rates are detected when comparing PSP with severe gaze palsy than less severe gaze palsy to healthy controls. The increased atrophy rates are detected in parts of the posterior frontal cortex involved in the generation of saccadic eye movements. Damage to these regions results in difficulty with intentional saccades and the antisaccade task, resulting in hypometric saccades in all directions. Hence, it seems that greater atrophy rates as a result of more severe pathological involvement of the cortical regions involved in the generation of saccadic eye movements, in combination with brainstem atrophy may result in the more severe gaze palsy in some PSP patients.

As cortical atrophy rates were greater in those with more severe gaze palsy, and in those with a higher FAB score, it is interesting to note that there was no significant difference in the mean FAB scores between the two groups.

Atrophy rates and the MMSE

Dividing the PSP patients according to the MMSE and comparing them with healthy controls revealed little difference in atrophy rates between the two groups when compared with controls and each other. Both groups demonstrated similarly increased brainstem atrophy rates compared to healthy controls. This suggests that the MMSE does not identify phenotypic differences with the PSP group. This is not surprising given the particular weightings (verbal) of the MMSE. It also argues for a consistent (neither accelerating nor decelerating) global loss.

Overall conclusions

The results of the fluid-SPM analysis demonstrate differing regional atrophy rates in PSP and MSA-P compared to healthy controls. Differing regional rates of atrophy were also clear when dividing the PSP group according to the FAB, disease duration and the severity of the gaze palsy. Interestingly, some of these regional atrophy rates were not identified when comparing the entire PSP cohort with healthy control subjects. This suggests that the PSP phenotype is varied and that this variation may arise as a consequence of differing rates of cerebral atrophy secondary to the distribution of pathology.

The studies of pathologically proven cases of PSP outlined in Chapter 3 of this thesis suggest that a common presentation of PSP may occur with a less severe, or absent gaze palsy, later falls (after one year), prominent bradykinesia, a flexed posture and a longer disease duration. In the PSP patient groups with a more severe gaze palsy and shorter disease duration, there was an overlap, with 5 of 8 patients being common to both groups. Those patients with a more severe gaze palsy, also had higher scores on the Hoehn and Yahr ($p=0.01$) and the UPDRS II ($p=0.03$), with a trend towards a higher score on the UPDRS III ($p=0.07$). There were no significant differences in the severity of motor deficit in the PSP patients separated on the basis of the FAB and the disease duration.

If we consider that short disease duration and severe gaze palsy are key to the discrimination of Richardson's syndrome from PSP-P then it is possible to conclude that those patients with these symptoms have faster rates of cortical atrophy in addition to the brainstem atrophy that is common to both subtypes. This may help to explain the faster rates of clinical disease progression seen in classical Richardsons

syndrome. Further prospective post mortem studies would be required to confirm this but in the meantime this information has important implications if using serially registered MRI and calculated atrophy rates to monitor disease progression. Further clinico-radiological correlates on the basis of ROI measurements are described in Chapter 12.

12. Correlating Clinical Features and Neuropsychology with Imaging.

12.1. Introduction

Over time, motor handicap and functional disability increase, in both PSP and MSA-P and in PSP particularly, there is also a decline in cognitive function (chapter 3.4.2). Data from the fluid–SPM studies in Chapter 11 support the theory that pathological involvement of specific brain regions results in the onset and evolution of specific clinical signs. An association between TIV corrected regional brain volumes with measures of disease severity would help to confirm this. Post mortem studies can infer this relationship is present when, for example a specific brain region is seen to be more severely affected by the pathological process in a group of patients with a specific clinical sign, but absent in a matched control group of patients. For example, in pathologically proven PSP without a clear history of a supranuclear gaze palsy (SNGP) during life, less severe pathological involvement of the nucleus raphe interpositus (omnipause neurons) (Revesz *et al.,* 1996), and the parietal eye fields (Verny *et al.,* 1996a) has been revealed. Pathological studies are often reliant on retrospective clinical data however, and only limited conclusions can be made. The considerable heterogeneity in the timing of the onset of the clinical features and in the rate of their progression, suggests that there is likely to be variation in the regional atrophy rates in PSP and MSA-P.

12.2. Cross sectional data

The objective of this section of the study was to determine whether regional atrophy, implied on the basis of cross sectional volumetric MRI analysis, was associated with the severity of the clinical features observed in PSP and MSA-P.

12.2.1. Methods

Scores on the UPDRS, Hoehn and Yahr rating scale, the MMSE, the FAB, the Schwab and England scale and the gaze palsy rating scale have been reported and compared in Chapter 4.

In this section, associations between ROI volume and motor deficit measured using the UPDRS and Hoehn and Yahr scores, as well as: falls; postural stability; speech and swallowing UPDRS sub-scores; the MMSE and; the FAB scores were assessed. In the PSP group, the severity of the gaze palsy was assessed using a scaled scoring system (Chapter 4) and its relationship with ROI volume evaluated. Any relationship between ROI volume and disease duration was also noted. The associations between the neuropsychological test scores (see Chapter 5) and ROI volumes in the PSP and MSA-P groups were also studied. Associations between frontal volume, and bedside and formal tests of executive function as well as associations between brainstem and cerebellar volumes and motor deficit were specifically studied.

All statistical analysis was undertaken using STATA version 8 (Stata Corp. College Station TX.). Linear regression was used to investigate the relationship between ROI and age in the healthy control group alone and to establish the relationship between TIV corrected ROI volumes, and neuropsychological and clinical information in the PSP and MSA-P groups. Clinico-radiological correlations in the PSP and MSA-P groups were studied separately using multiple regression analysis with age and disease duration as covariates.

The partial correlation coefficient was calculated where significant associations were detected in order to estimate the correlation that would occur if age and disease duration were held constant.

12.2.2. Results

Correlations between age and TIV corrected ROI volumes were only seen in the healthy control group: midbrain ($R^2 = 0.47$, $p = 0.001$); frontal volume ($R^2 = 0.41$, $p = 0.004$); and PI volume ($R^2 = 0.24$, $p = 0.04$). The disease related losses are likely to obscure the effect of age in the small disease groups. There were no significant correlations between ROI volume and disease duration in PSP or MSA-P.

Associations between midbrain volume and clinical rating scales (UPDRS II and III, Hoehn and Yahr, gaze palsy and postural instability) were present in the PSP patients (table 12.1). The midbrain volume was most strongly associated with motor deficit and postural stability as well as swallowing and gaze palsy. There were fewer significant associations between ROI volume and neuropsychological test scores in the MSA-P patients. The only significant associations between ROI volume and clinical scores in the MSA-P group were between: PI volume and FAB ($p=0.03$); cerebellar as well as pontine volume and FAB ($p=0.001$); and the third ventricle volume and swallowing ($p=0.04$) (table 12.2).

Table 12.1 *ROI correlations with clinical scores in PSP*

	Brain region																	
	whole brain volume		total frontal volume		PI volume		cerebellar volume		midbrain volume		pons volume		SCP volume		LV volume		third ventricle volume	
	Corr	p	Corr	P	Corr	p	Corr	P	Corr	p	Corr	p	Corr	p	Corr	p	Corr	p
MMSE																		
FAB															**-0.5**	**0.05**		
UPDRS2	**-0.6**	**0.02**	**-0.6**	**0.04**					**-0.6**	**0.02**								
UPDRS3	**-0.6**	**0.04**	**-0.6**	**0.04**					**-0.6**	**0.03**								
HY	**-0.6**	**0.03**	**-0.6**	**0.04**	*-0.5*	*0.07*			**-0.8**	**.001**			**-0.6**	**0.03**				
Falls	*-0.5*	*0.08*	**-0.6**	**0.02**	*-0.5*	*0.08*			*-0.5*	*0.06*			**-0.7**	**0.01**				
Postural Stability	**-0.6**	**0.02**	**-0.6**	**0.05**	**-0.6**	**0.05**			**-0.6**	**0.02**			*-0.5*	*0.06*				
Speech			*-0.5*	*0.06*									*-0.5*	*0.08*				
Swallowing	*-0.5*	*0.07*							**-0.8**	**.001**								
Disease duration																		
Gaze palsy score			**-0.5**	**0.05**					**-0.6**	**0.03**							*-0.5*	*0.08*

Significant positive or negative associations are shown in bold. Corr- partial correlation coefficient, p- p value (significance level at p=<0.05). PI- postero-inferior region, SCP-superior cerebellar peduncle, LV- lateral ventricle. All regression calculations have age and disease duration as co-variables.

Table 12.2 *ROI correlations with clinical scores in MSA-P*

	Brain region																	
	whole brain volume		total frontal volume		PI volume		cerebellar volume		midbrain volume		pons volume		SCP volume		LV volume		third ventricle volume	
	Corr	p	Corr	P	Corr	p	Corr	P	Corr	p	Corr	p	Corr	p	Corr	p	Corr	p
MMSE																		
FAB			*0.7*	*0.07*	**0.8**	**0.03**	**0.96**	**.001**	**0.72**	**.007**	**0.94**	**.001**						
UPDRS2																		
UPDRS3																		
HY																		
Falls																		
Postural Stability																		
Speech																		
Swallowing																	**0.78**	**0.04**
Disease duration																		

Significant positive or negative associations are shown in bold. Corr- partial correlation coefficient p- p value (significance level at p=<0.05). PI- postero-inferior region, SCP-superior cerebellar peduncle, LV- lateral ventricle. All regression calculations have age and disease duration as co-variables.

In the PSP group, more associations were demonstrated between frontal volume and neuropsychological tests of executive function (Reitan TMT, WCST, verbal fluency), than between whole brain volume and neuropsychological tests (table 14.3). Very few associations between neuropsychological test scores and regional volumes in MSA-P were detected.

Table 12.3 *ROI and Neuropsychology correlations in PSP group*

	DRS ip	DRS c	voc	Simi	VIQ	RMF	RTMT A/B	WCST c/n/p	VF F/A/S/An/Alt
BV	**0.55** ***0.05***	**0.61** ***0.03***	-	**0.62** ***0.03***	-	-	-	-	-
Frontal	**0.69** ***0.008***	**0.55** ***0.05***	**0.62** ***0.03***	**0.66** ***0.02***	**0.76** ***0.004***	**0.79** ***0.006***	^{A}**-0.71** ***0.03***	^{n}**-0.69** ***0.03***	^{an}**0.67** ***0.01*** ^{alt}**0.64** ***0.02***

BV-brain volume, Frontal-frontal volume. Upper number-partial correlation coefficient (assuming age and disease duration are constant), number in brackets-p value from regression equation with age and disease duration as co-variables.
DRS- Mattis dementia rating scale, ip-initiation, c-conceptualisation, voc-vocabulary (WAIS-R), simi-similarities (WAIS-R), RMF- recognition memory test faces, RTMT-Reitan trail making test (A and/or B), WCST-Wisconsin card sorting test,c-categories,n-errors,p-perseverative errors, VF- verbal fluency,letters F,A,S, An-animals, Alt-alternating semantic.

12.2.3. Conclusions

In PSP, the correlation between midbrain volume and clinical scales of disease severity and gaze palsy severity suggest a strong association between midbrain tissue loss and the evolution of the characteristic clinical signs and symptoms. All the patients fulfilled the diagnostic criteria for PSP, but no relationships between disease duration and ROI volumes were apparent. Disease duration is either error prone or there are very different rates of loss in different PSP subgroups. This is in keeping with the results of the fluid-SPM studies. Despite support from logistic regression analysis, some caution in interpreting the relationship between disease duration and clinical features or regional volumes is needed, as reporting of disease onset may not always be precise.

PSP may have different individual rates of progression, which together with ROI measurement error may contribute to the lack of a relationship between ROI volume and disease duration.

Increasing disease severity measured on UPDRS and Hoehn and Yahr scales, as well as falls and stability sub-scores, were significantly associated with a smaller midbrain volume, in-keeping with the results of previous studies (Groschel *et al.*, 2004). However, one previous VBM study failed to demonstrate this association and instead found a correlation between motor deficit (UPDRS III) and size of the caudate nucleus and motor cingulate cortex (Cordato *et al.*, 2005). Although these regions were not specifically studied here, atrophy of the motor cingulate cortex may contribute to the association of increased motor deficit, with reduced frontal volume.

Tissue loss in other regions may also contribute to the motor disability observed in PSP. For example, SCP volume was associated with Hoehn and Yahr score, with falls as scored on the UPDRS III, and there was a trend towards an association with

postural stability (table 12.1). In addition to ascending pathways (Chapter 6.3), some fibres run inferiorly to the brainstem reticular and vestibular nuclei and may be involved in balance and posture, helping to explain this association.

In the PSP group, reduced frontal volume was associated with increased motor disability (as already discussed) a more severe gaze palsy, falls and postural instability. Pathology in the midbrain and pons, particularly the nucleus raphe interpositus, has been suggested as responsible for the gaze palsy in PSP (Revesz *et al.,* 1996). The association between a smaller frontal volume and a more severe gaze palsy score could be explained by the contribution of pathology in the frontal eye fields to abnormalities of gaze. These results from the ROI data reassuringly support the results from fluid-SPM analysis. Motor impulsivity and reduced risk awareness may be partially responsible for increased falls in PSP and could help to explain the association between falls and reduced frontal volume However any association between frontal volume and gaze palsy could arise because frontal volume loss occurs in parallel with midbrain volume loss.

A lower frontal volume in PSP was also associated with poorer performance on tests of executive function including an increased number of errors on the WCST, prolongation of the RTMT, reduced verbal fluency and poor performance on subsets of the Mattis DRS dependent on frontal function. These associations were not seen with whole brain volume but remained when adding whole brain volume as a covariate into the regression model. Surprisingly, despite significantly lower FAB scores in PSP, and correlation between FAB scores and other tests of frontal cognitive function, there was no association between FAB score and frontal volume. However, caution must be exercised in drawing too many conclusions from these associations

because the template method for measurement of frontal volume also includes anterior temporal lobe structures and excludes some posterior frontal lobe structures.

In MSA-P, the only association between frontal volume and individual cognitive tests was with verbal fluency. There was no significant difference between frontal volumes in MSA-P and PD or controls. There was however an association between pons volume and FAB suggesting that in MSA-P subcortical pathology contributes to the modest executive dysfunction. There was also an association between cerebellar volume and the FAB as well as the RTMT, suggesting that cerebellar pathology may contribute to the pattern of cognitive dysfunction seen in MSA-P. An increased motor deficit seen in the MSA-P group was associated with smaller pontine volume.

The association of swallowing with third ventricle volume in MSA-P may reflect tissue loss in regions distant from this brain/CSF boundary, however functional imaging studies have concluded that structures surrounding the third ventricle are involved in swallowing (Harris *et al.*, 2005), suggesting a basis for the association seen.

Falls occur less frequently in MSA-P than PSP and often a little later in the course of the illness. Mean midbrain volume in MSA-P was only significantly lower than controls (not PD or PSP) and was not associated with increased motor disability scores. There was no association between any region and falls or postural stability in MSA-P. It may be that more severe pathological involvement of midbrain structures in PSP explains the early and severe postural instability.

12.3. Serial Imaging data

12.3.1. Introduction

If the pathological involvement and degeneration of a particular brain region results in the evolution of particular features of a disease, then the regional atrophy rate should be associated with an increase in the measured severity of those clinical features. This is of course dependent on the methods used to measure atrophy rates and evaluation of the change in the clinical features of the disease being reliable and sensitive (i.e. not having floor and ceiling effects).

12.3.2. Methods

In all patients with serial imaging, the change in scores on the Unified Parkinson's disease rating scale (UPDRS) and the Hoehn and Yahr scale, were recorded. The scores at time point one and time point two of: falls; postural instability; speech and; swallowing difficulty as recorded in the UPDRS were also noted. Changes in the Folstein MMSE and the frontal assessment battery (FAB) scores were also recorded in PSP and MSA-P. In PSP, the change in the gaze palsy scale scores were recorded. Where possible, patients completed a full neuropsychological assessment (chapter 5).

Annualised brain and regional atrophy rates were calculated by taking the BBSI and the regional BSIs in each group, adjusting for scan interval and expressing this as a percentage change from the baseline volume (Chapter 9).

Data were analysed using STATA version 8 (Stata Corporation, College station, TX).

For normally distributed data, linear regression was used to investigate the relationship between brain or regional atrophy rates and the change in clinical and neuropsychological test scores over time. For this analysis, age, disease duration and the baseline clinical test score were used as covariates. Partial correlation coefficients

(Corr.) were calculated to estimate the strength of the relationship if age and disease duration were constant. For non-normally distributed data, linear regression models with 95% bootstrapped confidence intervals (and implied p-values) were used to investigate the relationship between atrophy rates and change in test scores, with age, disease duration and the baseline clinical test as covariates.

In order to investigate whether regions were independently associated with change in the clinical scores of motor deficit (UPDRS and Hoehn and Yahr) and bedside tests of cognition (FAB and MMSE), multiple regression models were performed with regional atrophy rate as the dependent variable and change in clinical score as the independent variable. Age, disease duration, baseline clinical rating scale score and the next most significant regional atrophy rate were included as covariates.

12.3.3. Results

PSP

An increased rate of midbrain atrophy in PSP was associated with: a decline in the Mattis DRS-attention sub-test (Corr. =0.80, p=0.009); an increasing number of errors (but not perseverative errors) on the WCST (Corr. =0.72, p=0.04); a decrease in the PASAT score (Corr. =0.89, p= 0.04); a decrease in FAB score (Corr. =0.70, p= 0.008); an increase in motor disability (UPDRS III) score (Corr. =0.58, p=0.04) and a decrease in the MMSE (Corr. =0.56, p=0.04). Increased pontine atrophy rates were only associated with a decrease in FAB score ($p<0.05$). Increasing cerebellar atrophy rates were associated with decrease in the FAB score (Corr. =0.73, p=0.005) and in the MMSE score (Corr. =0.55, p=0.05).

Increased whole brain atrophy rates were associated with decline in performance on: the Mattis DRS attention sub-test (Corr. =0.77, p=0.009); an increasing number of

errors on the WCST (Corr. =0.83, p=0.006); an increase in the number of intrusions on the verbal fluency test (Corr. =0.89, p=0.003); a decline in performance on the PASAT (Corr. =0.82, p=0.045); and a reducing score on the FAB (Corr. =0.65, p=0.02) and the MMSE (Corr. =0.68, p=0.01).

Increased rates of frontal and postero-inferior atrophy were similarly associated with a more severe deterioration in: the Mattis DRS attention sub-test (frontal, Corr. =0.70, p=0.02; postero-inferior, Corr. =0.65, p=0.04); the number of errors on the WCST (frontal, Corr. =0.85, p=0.004; postero-inferior, Corr. =0.72, p=0.03); a decline in the PASAT score (frontal, Corr. =0.94, p=0.005; postero-inferior, Corr. =0.83, p=0.04); the FAB (frontal, Corr. =0.62, p=0.02; postero-inferior, Corr. =0.67, p=0.01); the MMSE (frontal, Corr. =0.56, p=0.05; postero-inferior, Corr. =0.75, p=0.003). An increasing rate of postero-inferior atrophy was also associated with increased motor disability measured on the UPDRS III (Corr. =0.56, p=0.05).

An increasing rate of third ventricle enlargement was associated with a decline in performance on: the DRS attention subtest (Corr. =0.73, p=0.02,); the number of errors on the WCST (Corr. =0.80, p=0.009); an increasing number of intrusions on the verbal fluency test (Corr. =0.81, p=0.014) and a decline on the PASAT (Corr. =0.93, p=0.008). An increasing rate of lateral ventricle enlargement was only associated with a decline in the FAB ($p<0.05$) and a decline in the MMSE ($p<0.05$).

Having adjusted for age and disease duration, decline in the FAB score was most strongly associated with the rate of midbrain atrophy (p=0.008). The next strongest association was with atrophy rate in the cerebellum (p=0.005), but neither of these had independent effects once the other was controlled for. Change in UPDRS III was most strongly associated with midbrain atrophy (p=0.04) and then posterior inferior

atrophy. Again, neither of these had independent effects once the other was controlled for. Similar results were obtained for change in the MMSE.

MSA-P

In the MSA-P patients, increased rates of pontine atrophy were associated with a decline in performance on: the DRS total score (Corr. =0.99, p=0.001); the digit span (Corr. =0.88, p=0.05) and semantic verbal fluency (Corr. =0.95, p=0.05) as well as an increase in the motor disability as scored on the UPDRS III (Corr. =0.8, p=0.05).

Increasing cerebellar atrophy rates were also associated with declining performance on the DRS total score (Corr. =0.96, p=0.01); the DRS conceptualisation subtest (Corr. =0.89, p=0.04); semantic verbal fluency (Corr. =0.95, p=0.05) and UPDRS III (Corr. =0.86, p=0.03).

Increasing midbrain atrophy rates were associated with poorer performance on: the DRS initiation and perseveration subtest (Corr.=0.97, p=0.005); the recognition memory test for faces (Corr.=0.92, p=0.03); the digit span (Corr. =0.91, p=0.03) and intrusions on the verbal fluency test (Corr.=0.96, p=0.04).

Lower whole brain, frontal and postero-inferior atrophy rates were associated with a greater rate of decline in the PASAT score ($p<0.05$) and for frontal atrophy rates, the Beck anxiety inventory (p=0.04).

In the control subjects, an association, albeit weak, between age and increasing rates of lateral ventricle enlargement (p=0.04) and posterior-inferior atrophy (p=0.04) were noted. No associations between age or disease duration and atrophy rates were seen in PSP or MSA-P.

As with the associations noted in PSP, multiple regression analysis confirmed that change in motor deficit was most strongly associated with cerebellar and pontine atrophy rates but that neither of these had independent effects.

12.3.4. Conclusions

There are a number of associations between atrophy rates and the clinical features of these diseases. Crucially, rates of atrophy in PSP and MSA-P have clinical correlates reflecting progression in clinically meaningful end-points. Worsening motor deficit was associated with increasing midbrain atrophy rates in PSP and increased ponto-cerebellar atrophy in MSA-P. This suggests that pathological involvement of the structures in these regions, represented by tissue loss (and presumably cell death), is strongly associated with progression in the critical clinical features of the diseases. Hence regional MRI derived atrophy rates, which correlate with change in motor deficit in PSP and MSA-P, may be useful and quantifiable outcome measures in clinical trials. This is evaluated further in Chapter 14.

Associations between atrophy rates and change in neuropsychological and clinical scores were largely confined to subtests of the Mattis DRS, the WCST and the PASAT. Associations between increasing rate of frontal atrophy in PSP and an increasing number of errors on the WCST, and declining PASAT and FAB scores, suggests that the degree of severity noted in the dysexecutive syndrome in PSP is related to frontal atrophy. This supports observations from cross sectional imaging studies (Cordato *et al.*, 2002; Cordato *et al.*, 2005) and also supports the findings of the fluid-SPM study, which showed that the greatest rates of frontal cortical atrophy occur in patients with the highest FAB scores. Damage to brainstem and sub-cortical structures also contributes to the cognitive dysfunction in PSP and hence it is not

surprising that there is a strong association between a decline in the FAB score and increased rates of midbrain atrophy in PSP.

The association between a greater decline in the FAB score in PSP and increasing lateral ventricle atrophy rates could in part, be explained by changes at the lateral ventricle/caudate nucleus boundary. Atrophy of the caudate nucleus in PSP could contribute to a worsening sub-cortical dementia. However, changes in lateral ventricle volume (measured using the BSI) are unlikely to reflect tissue loss simply from contiguous structures and are more likely to arise as a consequence of distant tissue loss and subsequent structural readjustments.

Correlations were also seen between some of the neuropsychological tests and whole brain rates of loss as well as posterior inferior atrophy rates in PSP. It is possible that the associations arise because the tests applied are those most sensitive to change over a relatively short period of time. Other tests may be less sensitive or may be at floor at the first assessment.

The smaller number of patients with MSA-P is also likely to have limited the associations detected in this group.

Additionally, the results of multiple regression models, using age, disease duration and the next most significantly associated regional atrophy rate as covariates suggest that some of the associations seen between regional atrophy rates and change in disease severity arise because atrophy in one region is likely to occur in parallel with another region. For example, the association between pontine atrophy and change in the FAB score in PSP may only arise because the real association is with midbrain

atrophy, and pontine atrophy and midbrain atrophy occur in parallel. In order to clarify this, larger study numbers would be required.

The results outlined in chapters 9-12 confirm that volumetric imaging has a potential role as a marker of disease progression, as measured atrophy rates have a clinical correlate, one of the criteria for a good surrogate marker. The results of cross sectional diffusion weighted MRI analysis suggest that it can be applied to detect regional pathological differences and destruction of normal tissue architecture, between PSP, MSA-P PD and healthy controls. Whether it too has a potential role, as an imaging marker of disease progression over a short interval requires assessment.

13. Serial diffusion weighted imaging

13.1. Introduction

Diffusion weighted MRI (DWI), allows detection of damage to brain tissue, as reduced "anisotropy" results in an increased apparent diffusion coefficient (ADC, Chapter 8). In neurodegenerative diseases, progression of symptoms over time occurs secondary to increasingly severe pathology in specific regions of the brain. These pathological changes can be tracked by measuring changes in brain volume (Chapter 9). If the ADC value of a tissue is influenced by the degree of damage to the normal tissue architecture, then it would be reasonable to hypothesise that changes in the ADC value of a region over time may be used to track the progress of a disease. Whether measuring the difference in rADC over time (less than 1 year) is sensitive enough to measure this change is evaluated in this section.

13.2. Methods

Follow up scans were conducted using the same DWI protocol as used for the baseline scans (Chapter 8). Image analysis was undertaken in the same way. As before, the rater was blinded to the identity and diagnosis of the subject and the images presented in a random sequence. ADC values were recorded for all 18 regions (Left and right middle cerebellar peduncle-MCP, caudal and rostral pons, midbrain, decussation of the superior cerebellar peduncle-SCP, left and right thalamus, left and right caudate nuclei, left and right putamen and globus pallidus, genu and splenium of the corpus callosum, the left and right frontal and parietal white matter and the white matter of the left and right centrum semi-ovale). The annualised difference between the baseline and follow up analysis was calculated by subtracting the baseline ADC value from the follow up ADC value and correcting the result for an annual interval.

Non-parametric data were analysed using a Kruskal-Wallis test, with post hoc Mann Whitney U tests where significant group differences were observed. Normally distributed data were analysed using a one-way analysis of variance (ANOVA) with post hoc t-tests where significant group differences were identified.

In addition, the mean value from combined left and right regions was determined and the annualised difference between baseline and follow up values calculated.

13.3. Results

In total, 16 PSP, 9 MSA-P, 9 PD patients and 6 healthy control subjects had serial DWI, and regional ADC measurements. The mean (SD) scan interval (days) was not significantly different between the groups (PSP- 288.8 (59.7), MSA-P-262.3 (65.5), PD-269.2 (48.7), Controls-217.8 (48.7), p=0.1).

The annualised changes in rADC values are detailed in the table below.

Table 13.1 Mean (SD) rADC difference ($x10^{-3}mm^2s^{-1}$) per annum.

	PSP	MSA-P	PD	HC	P values
N	16	9	9	6	
L MCP	0.038 (0.14)	-0.036 (0.09)	0.008 (0.05)	0.016 (0.08)	NS
R MCP	0.056 (0.22)	-0.032 (0.14)	-0.006 (0.09)	-0.058 (0.12)	NS
Caudal Pons	0.050 (0.21)	0.073 (0.12)	0.060 (0.29)	0 (0.24)	NS
Rostral Pons	-0.001 (0.17)	-0.023 (0.16)	-0.001 (0.06)	-0.118 (0.23)	NS
Midbrain	0.031 (0.12)	-0.004 (0.12)	0.027 (0.10)	-0.078 (0.12)	NS
SCP	0.023 (0.19)	-0.017 (0.11)	0.011 (0.14)	-0.051 (0.098)	NS
L Thalamus	0.050 (0.14)	0.021 (0.05)	0.021 (0.08)	0.017 (0.07)	NS
R Thalamus	0.046 (0.16)	-0.043 (0.07)	-0.009 (0.08)	-0.026 (0.08)	NS
L putamen	0.031 (0.19)	0.018 (0.06)	0.086 (0.15)	-0.021 (0.04)	NS
R putamen	0.016 (0.15)	-0.003 (0.07)	0.041 (0.12)	-0.044 (0.06)	NS
L Globus pallidus	0.057 (0.21)	0.019 (0.08)	0.048 (0.13)	-0.01 (0.07)	NS
R Globus pallidus	0.038 (0.15)	0.012 (0.07)	0.043 (0.15)	-0.003 (0.04)	NS
Splenium CC	0.029 (0.26)	0.142 (0.22)	0.019 (0.31)	-0.337 (0.270)	0.01*
Genu CC	0.100 (0.28)	-0.156 (0.19)	0.090 (0.34)	-0.035 (0.198)	NS
L Caudate	0.074 (0.19)	-0.018 (0.14)	0.020 (0.086)	-0.087 (0.097)	NS
R Caudate	0.045 (0.21)	-0.076 (0.19)	0.004 (0.074)	-0.080 (0.131)	NS
L Frontal WM	0.075 (0.21)	-0.027 (0.06)	0.016 (0.026)	0.011 (0.059)	NS
R Frontal WM	0.046 (0.15)	-0.008 (0.04)	0.013 (0.034)	-0.019 (0.045)	NS
L Parietal WM	-0.018 (0.11)	-0.040 (0.06)	0.008 (0.048)	-0.017 (0.050)	NS
R Parietal WM	0.020 (0.09)	-0.049 (0.07)	-0.001 (0.07)	-0.069 (0.026)	0.05#
L Centrum Semi Ovale	0.036 (0.13)	0.007 (0.05)	0.028 (0.069)	-0.026 (0.017)	NS
R Centrum Semi Ovale	0.025 (0.14)	0.016 (0.06)	0.035 (0.079)	-0.027 (0.049)	NS

Post hoc tests: * PSP vs. HC, p=0.006, MSA vs. HC, p=0.005; # PSP vs. HC, p=0.01

Table 13.2 Mean (SD) combined left and right rADC difference ($x10^{-3}mm^2 s^{-1}$) per annum.

	PSP	MSA-P	PD	HC	P values
N	16	9	9	6	
MCP	0.047 (0.18)	-0.035 (0.11)	0.001 (0.06)	-0.021 (0.09)	0.6
Thalamus	0.048 (0.14)	-0.011 (0.05)	0.006 (0.07)	-0.004 (0.07)	0.8
Putamen	0.024 (0.17)	0.007 (0.06)	0.063 (0.13)	-0.032 (0.02)	0.09
Globus Pallidus	0.048 (0.17)	0.015 (0.07)	0.045 (0.14)	-0.006 (0.05)	0.9
Caudate	0.062 (0.19)	-0.06 (0.19)	0.012 (0.07)	-0.084 (0.10)	0.1
Frontal WM	0.061 (0.18)	-0.017 (0.05)	0.015 (0.02)	-0.004 (0.05)	0.4
Parietal WM	0.001 (0.09)	-0.044 (0.06)	0.004 (0.06)	-0.043 (0.02)	0.2
Centrum Semi Ovale	0.031 (0.13)	0.011 (0.03)	0.032 (0.07)	-0.026 (0.03)	0.2

The right parietal white matter and the splenium of the corpus callosum were the only regions in which significant differences in the change in ADC value over time between the groups were found. When the left and right regions were combined and differences between these mean values from baseline to follow up were analysed, no significant differences were identified (table 13.2).

13.4. Conclusions

The results suggest that diffusivity measured using the regional ADC is not sensitive to change over time in PSP, MSA-P or PD. From this we can conclude that DWI seems unsuitable for measuring disease progression in these conditions, over a short interval. The significant differences identified in the change in rADC values for the splenium and the parietal white matter arose because of an apparently negative (i.e a decrease in the rADC) difference between baseline and follow up values in the control population, rather than because of an increase in the rADC value in any of the patient groups. Measurement error is likely to have influenced the results and any real difference in the baseline and follow up rADC values, is likely to be obscured by this. In addition, while in the MCP, the pons, midbrain, thalamus, and the white matter regions the repeatability and reliability (within subject SD and the ICC) of the

measurements were good, in the SCP, the caudate nucleus, the putamen and globus pallidus, repeatability and reliability of the measurements was poor.

The only previous study of longitudinal DWI in MSA-P detected a significant difference between 8 MSA-P and 8 PD patients, in the putamen (Seppi *et al.*, 2004; Seppi *et al.*, 2005) had a longer mean scan interval of 15 months. The ADC value in the putamen has been shown to discriminate MSA-P from PD and healthy controls in cross sectional studies (Schocke *et al.*, 2002). In the cross sectional data in this study, the MCP and pons had significantly greater ADC values in MSA-P compared with PSP, PD and healthy controls (Chapter 8). Hence it is perhaps surprising that no difference between the change in the pons rADC value in MSA and healthy controls was detected. Several factors could contribute to this failure. Firstly, DWI may not be sensitive enough to detect the mild changes occurring over the time interval in this study. Secondly, failure to measure exactly the same ROI on the follow up ADC map could create inaccurate results and contribute to the large variation in the mean rADC differences seen in most of the regions measured. Unlike the volumetric image analysis, DWI scans were not registered to standard space. This means that it is unlikely the exact same anatomical regional ADC was taken for the baseline and follow up scan. This is difficult to correct as registering the DWI images to standard space would result in a loss of anatomical resolution and subsequently new problems regarding the identification of regions of interest.

Measurement error secondary to partial volume artefact was minimised as pixels containing predominantly CSF were excluded from the analysis by means of thresholding (Chapter 8).

It seems that the methods of serial DWI cannot be applied as a measure of disease progression. This is in contrast to the conclusions of serial volumetric imaging

studies. However, in order to confirm that volumetric imaging has a role in monitoring disease progression, it needs to be demonstrated that a feasible number of patients enrolled in a drug trial will allow detection of that drug effect. This can be assessed using a power calculation.

14. Power calculations using BBSI and regional BSI

14.1. Sample size calculations

14.1.1. Introduction

Advances in the understanding of the pathological processes that underlie PSP and MSA means that potential disease-modifying drugs may soon become available. The disease modifying effect of these drugs will need to be quantified. Measurement of a true disease modifying effect with clinical scales can be difficult and it would be helpful to be able to apply imaging as a truly quantifiable marker of disease progression. Atrophy rates in specific regions in PSP and MSA-P have been demonstrated to have clinically meaningful correlates suggesting potential for serial volumetric MRI as a marker of diseases progression. However, determining the feasibility of atrophy rates as a marker of disease progression depends upon the sample size required to detect the slowed atrophy rate (the drug effect).

An essential part of planning any drug trial is an estimation of sample size required to show the desired effect of a drug. Justifying the proposed study size and demonstrating that the study is capable of answering the questions posed is an important component of any research protocol.

If the effect of a drug is to be measured using a quantifiable marker of disease progression, then one has to decide what percentage difference in that marker between diseased patients treated with placebo and those treated with the drug, one would want to be able to detect (for example 20% slowing) (Fox *et al.*, 2000). For a thorough power calculation, the significance level (or P-value) required in order to reject the null hypothesis of no difference in rates of atrophy, between a treated and untreated group needs to be stated. Finally the probability that we would like to have of

achieving this level of significance needs to be stated. This is required since because of sampling variation, the possibility that the size of the effect observed in the study will be smaller than the true effect cannot be ruled out. Increasing the sample size increases the power of the study and the probability of detecting a significant effect.

The pathological process that underlies PSP and MSA-P is greatest in subcortical regions in the early stages of the disease. Although cortical pathology is known to occur as both diseases progress (Bigio *et al.,* 1999; Konagaya *et al.*, 1999a; Verny *et al.,* 1996a) and the fluid-SPM studies conducted as part of this thesis suggest cortical atrophy may occur early in PSP, using whole brain atrophy rates as a marker of progression in PSP and MSA-P may not be practical. Compared to Alzheimer's disease, the number of subjects required in each arm of a drug trial to generate sufficient power to detect a disease modifying effect may be too large because rates of whole brain atrophy are only increased by a small amount in the early stages of the disease. The cortical atrophy that occurs in PSP is relatively focal. In PSP the midbrain is known to be markedly affected by the disease process and in MSA the pons and cerebellum are involved (Ozawa *et al.,* 2004) resulting in macroscopic atrophy at post mortem. Fewer subjects may need to be recruited to a study to generate statistical power if regions in which atrophy rates are greatest are used as a marker of disease progression in PSP and MSA-P.

The aim of this section of the study was to determine the feasibility of using whole brain and a number of hypothesis driven regional atrophy rates in PSP and MSA-P as measures of disease progression in clinical trials. The number of individuals needed to power a clinical trial of a proposed disease-modifying agent was used as a measure of feasibility.

14.1.2. Methods

Whole brain BSI (BBSI) and regional BSIs (Pons, midbrain, frontal quadrants, lateral ventricle and third ventricle) measures of atrophy rates were assessed and sample size estimates (power calculations) performed for potential disease modification trials. Separate power calculations were made for PSP and for MSA-P.

Data were analyzed using Stata version 8.0 (Stata Corporation, College Station, Tex, 1997). Sample size requirements were estimated for whole brain and regional rates of atrophy expressed as a percentage of initial brain or region volume and annualised to give a yearly rate of atrophy. The following standard formula was applied to generate sample sizes (Kirkwood and Sterne, 2003) in the PSP and the MSA-P patients:

$$\text{Sample size} = (u + v)^2 \times (\sigma_1^2 + \sigma_2^2) / (\mu_1 - \mu_2)^2$$

Where: u = 1.28 (the one sided percentage point of the normal distribution corresponding to 100%- the power) to provide 90% power and v = 1.96 (the percentage point of the normal distribution corresponding to the two-sided significance level) to test at the 5% level; μ and σ are the mean and SDs of rates of atrophy (measured using the whole brain BSI and the regional BSIs); σ_1- standard deviation in the placebo group (taken as the measured sd from the PSP, MSA or PD groups); σ_2- standard deviation in the treatment group, presumed to equal that of the placebo group; μ_1-mean BSI of the placebo group (taken as the measured mean BSI from the PSP or MSA-P groups); μ_2- mean BSI in the treatment group (estimated as a 20%, 30%, 40% or 50% reduction in the rate of atrophy).

The following modifications were included:

1. The difference in annual rates of tissue loss between patients with PSP or MSA-P and healthy age-matched controls was taken to represent the maximum possible effect a treatment could have (ie, 100% impact). If this assumption is not made, the potential for a therapeutic effect may be overestimated. A reduction of 20% in progression was therefore considered to be equal to 20% of this difference rather than 20% of the total loss in patients with PSP or MSA-P.
2. The sample size thus derived was increased by 10% to allow for losses to follow-up.
3. The sample size was then further increased assuming that about 10% of scan pairs could not be used (e.g., too much movement on one of the scans).

Hence the adjustment factor for 20% loss = 100/ (100-20) = 1.25, which approximates to the number of individuals who after recruitment to the study did not have a complete imaging dataset. All calculations were performed based on the requirement that a trial should have 90% power to detect the specified treatment effect when a 2-sided 5% level of significance is used ($p<0.05$).

14.1.3. Results

The healthy control group and the PSP and MSA-P groups were closely matched for age and sex. The atrophy rates expressed as percentage tissue lost per annum (reported in Chapter 9) were used in the power calculations.

The results of the power calculations can be interpreted as follows: For a drug with an anticipated ability to reduce the rate of cerebral atrophy in PSP by 20%, over 1 year, to have 90% power to detect a drug effect (allowing for atrophy due to normal aging, a 10% dropout rate, and assuming that 10% of scan pairs are unusable) 1122 patients

will be needed in each treatment arm. If the drug reduces atrophy rates by 50%, only 179 patients would be required per treatment arm (table 14.1). Table 14.2 shows the number needed to power studies of MSA-P patients for varying rates of atrophy.

Table 14.1 *Necessary Sample Size per Treatment Arm Using Group Mean Rates of Atrophy*

PSP									
% reduction in atrophy rate)	Number of individuals required in each arm of the trial								
	BBSI	Pons BSI	Midbrain BSI	Lateral ventricle BSI	Third ventricle BSI	SCP BSI	Cereb BSI	Frontal BSI	Posterior Inferior BSI
20	1122	623	412	1559	1516	853	796	1691	1278
30	499	277	183	693	674	379	354	752	568
40	281	156	103	390	379	213	199	423	319
50	179	100	66	249	243	136	127	271	204

BBSI-brain boundary shift integral, BSI-boundary shift integral, SCP- superior cerebellar peduncle, Cereb-Cerebellum

Table 14.2 *Necessary Sample Size per Treatment Arm Using Group Mean Rates of Atrophy*

MSA-P									
% reduction in atrophy rate)	Number of individuals required in each arm of the trial								
	BBSI	Pons BSI	Midbrain BSI	Lateral ventricle BSI	Third ventricle BSI	SCP BSI	Cereb BSI	Frontal BSI	Posterior Inferior BSI
20	1786	369	1052	2596	2340	4852	290	3620	663
30	794	164	467	1154	1040	2156	129	1609	295
40	447	92	263	649	585	1213	72	905	166
50	281	59	168	415	374	776	46	579	106

BBSI-brain boundary shift integral, BSI-boundary shift integral, SCP- superior cerebellar peduncle, Cereb-Cerebellum

Calculations based on regional rather than whole brain rates of atrophy confirm that fewer subjects would be required to power a trial of a disease modifying drug.

14.2. The effect of scan interval on variance

14.2.1. Introduction

The variance in the calculated atrophy rates arises mainly as a consequence of measurement error and true inter-individual variability in atrophy rates in patients with the same disease. This variability in atrophy rates is one of the factors that

influences the number of patients required to power a trial of a disease-modifying drug. If MRI is to be applied as a method of monitoring disease progression and quantifying the effect of disease modifying drugs, then determining how the scan interval influences variability is important.

14.2.2. Methods

In order to investigate whether the calculated rate of atrophy and the variance in the measured atrophy rates is significantly influenced by the scan interval, the PSP patients were divided into those with a scan interval below the median (50^{th} centile) of 300 days and those above. Mean annual percentage atrophy rates calculated using the BSI were compared, using a t-test for normally distributed data and a Kruskall Wallis test for non-parametric data. A Pitman's test was used to determine whether increasing the scan interval significantly improved (i.e. reduced) the variance in atrophy rates.

14.2.3. Results

Nine PSP patients had MRI scans 300 days or more apart and 8 had scans 299 or fewer days apart. The mean age and disease duration in these two groups was not significantly different and the mean annualised percentage volume loss in the whole brain, midbrain, pons, cerebellum, SCP, frontal and posterior inferior regions and the percentage expansion in the ventricles were no different between the groups (table 14.3).

However the results of Pitman's test of the equality of variance in the two groups demonstrates that the variance in third ventricle and lateral ventricle measurements was significantly greater in the patients with a shorter scan interval. There was a similar but non-significant trend towards a greater variance in measured midbrain and pontine atrophy rates in patients with a shorter scan interval (table 14.3).

Table 14.3 *Comparing mean (SD) % atrophy /year, in PSP with different scan intervals.*

	PSP <300d	PSP 300d+	P value	Pitman's test
Brain	1.2(1.3)	1.1(0.8)	0.8	0.1
MB	2.5(1.9)	1.9(1.0)	0.4	0.06
Pons	1.8(1.6)	1.3(0.9)	0.8	0.07
Cerebellum	1.5(1.2)	1.3(1.1)	0.7	0.5
SCP	3.1(3.9)	3.8(4.2)	0.7	0.6
Frontal	1.5(1.5)	0.9(1.0)	0.3	0.2
PI	1.1(1.2)	0.9(0.8)	0.6	0.1
LV	11.5(13.7)	7.9(4.2)	0.6	0.002
3V	7.3(8.3)	5.1(2.9)	0.9	0.004

PSP <300d- PSP patients with scan interval less than 300 days, PSP 300d+ - PSP patients with scan interval of 300 days or more.

These results suggest that while a shorter scan interval may mean less drop-out, the number of patients required to power a trial is likely to be greater because of the increased variance in measured atrophy rates.

14.3. Overall Conclusions

The calculations presented above confirm that measuring regional rather than whole brain rates of atrophy allows for recruitment of fewer patients to clinical phase 2/3 drug trials in PSP. Recruiting several hundred subjects for each arm of a drug trial using whole brain atrophy rate as a surrogate marker would require a large, multi-centre international trial. This would be more expensive to run and to evaluate than a study utilising regional rates of atrophy that might only need a half or even one third of the patient numbers in each arm to power the trial.

Recruiting more subjects than necessary to power a study is not only more expensive, but exposes more patients to the risks and potential side effects (however minimal) of a novel therapy.

In PSP and MSA-P, sample size calculations based on brainstem atrophy rates are of the order of three to five times smaller than those based on whole brain atrophy rates. In PSP, sample sizes were lowest when using midbrain atrophy rates (183 patients per study arm to detect a 30% treatment effect). In MSA-P, Cerebellar atrophy rates generated the smallest required sample sizes (129 patients per study arm to detect a 30% treatment effect). This number takes into account patient drop out and problems with scan quality. Even fewer study numbers would be required if 100% attendance could be achieved or if the drug effect were greater. This confirms that regional atrophy rates rather than whole brain atrophy rates are more feasible markers of disease progression in PSP and MSA-P.

The calculations above take into account subject attrition and technical problems with scan acquisition, yet still indicate a potentially realistic sample size given the cost of including serial MRI in a large clinical trial. The therapeutic effects used in the examples given ranged from 20% to 50%. Less than a 20% disease-slowing effect would be difficult to show and might not be considered clinically relevant. A disease-slowing effect of more than 50% is unlikely to need a surrogate marker of progression because the clinical effects would be striking and obvious.

The factors that are important in determining sample size are: Firstly, the difference between PSP/MSA-P and control rates of atrophy for the structure measured; the maximum treatment effect that could be expected is to reduce the rate of loss to that of age-matched controls (100% disease slowing); the second factor is the heterogeneity

or variability (the standard deviation) of the measured rates of atrophy in the PSP or MSA group. This variability is the result of genuine differences between individuals, as well as the measurement error of the technique. The more homogeneous, the measured rate of atrophy, the easier it is to detect a small treatment change. The results in Chapter 14.2.3 suggest that a short scan interval will increase the variability of measured atrophy rates and is likely to increase the number of subjects required to generate statistical power.

Techniques that measure brain substructures may be subject to greater measurement error because of the difficulty in manually segmenting (outlining) these structures. The use of manual methods of segmentation and volume measurement introduces the issue of inter-observer and intra-observer variability. Measurement errors are likely to be greater for smaller and less easily defined neuroanatomical structures, particularly if an arbitrary cut-off to the structure is used (as in the case of the SCP in this study). In these structures, while the difference in atrophy rates between diseased patients and controls is great, the variability in the measurement is also high, meaning it is not practical to use such regions as markers of disease progression.

Brain substructures may have variable rates of loss as the disease progresses. For example, it may be that while the superior cerebellar peduncle atrophies early in the course of PSP, it may not lose tissue as rapidly in the later stages of the disease. This is also seen in AD where Jack *et al* (Jack, Jr. *et al.*, 1997) found that in very mild AD, the mean hippocampal volume was already significantly smaller than that of controls. More severely affected patients had a mean hippocampal volume that was only slightly smaller than that of the very mildly affected patients.

PSP and MSA-P are heterogeneous neurodegenerative conditions, and rates of disease progression between subjects are variable. Using subjects as their own controls for rates of atrophy might allow for this variability. Subjects would need to have a 3- to 6-month placebo run-in with 2 MRI scans to establish their individual rate of atrophy. A further MRI scan would then be performed after 9 to 12 months on either active treatment or placebo, with the outcome assessed by comparison with their baseline rate of atrophy. Unfortunately because of the increased number of scans required, the cost of running such a study would be greater and drop out rates might be higher than the predicted 10%.

Regional atrophy rates have been studied in other degenerative diseases but most commonly in AD. Ventricular atrophy rates show greater separation between cognitively normal subjects and AD patients than whole brain measures of volume change (Gunter *et al.,* 2003) and effect size and power calculations have suggested that in AD, entorhinal cortex and hippocampal measurements may be more sensitive than ventricular or whole brain BSI for detecting progression and the potential effects of disease modifying agents (Ezekiel *et al.*, 2004).

In PSP and MSA-P, both less common diseases than AD, reducing the number of patients needed to power a drug trial is important as the calculated number required using whole brain atrophy rates is prohibitively high.

15. Summary

Neurodegenerative diseases that result in bradykinetic rigid syndromes can be difficult to distinguish on clinical grounds alone. A non-invasive test that supports the diagnosis of PSP would have considerable clinical value. The advent of symptomatic therapies for PSP and in time, disease modifying drugs means that accurate diagnosis in the early stages of the disease will become even more important.

- Identifying an atrophic superior cerebellar peduncle (SCP) on cross sectional MRI supports a diagnosis of PSP over MSA-P or PD, and considering this as well as other regions together rather than independently increases its diagnostic usefulness as it is the pattern of loss rather than the appearance of a single structure of interest that is important.
- Serial volumetric MRI analysis is a feasible method of monitoring disease progression and would allow the effect of a putative drug that slowed atrophy rates by 20% to be detected.

The conclusions related to the specific aims of the study detailed in the opening section are outlined below.

MRI as an aid to diagnosis in PSP

The clinical utility of region of interest volume measurements in discriminating PSP from degenerative diseases that may present with a similar phenotype is supportive of the diagnosis when the clinical diagnostic criteria are fulfilled and may help to discriminate cases that cannot clearly be categorised on the clinical diagnostic criteria alone. Because volumetric analysis is time consuming, the direct clinical application of this technique may be limited, however the results of this thesis include new data, which can be applied in this context.

- The superior cerebellar peduncle (SCP) can be clearly visualised using MR imaging and atrophy in this region is highly specific for PSP (chapter 6). Serial imaging studies also demonstrate that the region is clearly more affected by progressive atrophy in PSP than in MSA-P or PD. An important conclusion is that the clinician should look at the SCP when seeking imaging features to support the clinical diagnosis of PSP.
- Midbrain volume alone is not a reliable discriminator of PSP from MSA-P (specificity of 33%) and statistical analysis of the discriminating ability of various regions of atrophy confirms that discrimination of PSP from MSA-P is better supported by considering a collection of imaging features together rather than independently. In all likelihood this is, subconsciously, what an experienced neuroradiologist does when interpreting an MRI scan. This study also identifies the regions that when assessed together may best allow identification of patients with PSP.

Diffusion weighted imaging provides information regarding tissue integrity as the apparent diffusion coefficient (ADC) is altered by any pathological process that affects the normal tissue architecture.

- Cross sectional diffusion weighted MRI demonstrated that mean regional apparent diffusion coefficients in the middle cerebellar peduncle were able to differentiate MSA-P from PSP, PD and healthy controls. This suggests that disruption to the normal tissue architecture as a consequence of the pathological process can be detected in life using MRI. Similarly DWI was able to demonstrate involvement of the frontal white matter in PSP. The results of serial DWI analysis suggest that over the relatively short interval in

this study (<1 year), disease progression cannot be monitored on the basis of change in regional ADC values.

DWI sequences are rapidly acquired and may still have utility in a cohort of patients who might not tolerate longer scan sequences. Changes to normal tissue architecture may be detected using DWI earlier than marked volume loss and allow discrimination of PSP from MSA-P earlier in the disease course. Further studies of DWI and ADC values in these diseases in their earliest stages would be interesting.

Novel image analysis techniques

The region of interest studies conducted, as part of this research, required that hypothesis driven a priori assumptions regarding the regions to study needed to be made.

- The results of unbiased assessments of cross sectional volumetric scans (VBM) confirm that these a priori assumptions were reasonable, with the most significant regions of volume difference being identified in the brainstem in PSP and MSA-P. The VBM studies also identified the middle cerebellar peduncle; a region not studied using manual segmentation, as discriminating MSA-P. This led to assessment of this region using diffusion weighted MRI.
- Fluid-SPM analysis, a way of assessing atrophy rates with no a priori assumptions regarding where that atrophy occurs, demonstrates that atrophy rates in PSP are greatest in the upper brainstem. Cortical atrophy occurs predominantly in anterior structures and occurs early in the course of the disease, contributing to the pattern of cognitive impairment typical in PSP.
- Fluid-SPM analysis also suggests that the frontal cortical atrophy may occur earlier in the disease course, however it may simply be that those patients with

a longer disease duration have slower rates of brain atrophy and clinical progression.

- Those PSP patients with more severe gaze palsy, have differing patterns of ongoing atrophy suggesting more severe involvement of the cortical regions responsible for the generation of saccadic eye movements in addition to the brainstem regions known to be affected. This has important implications if serial MRI is to be considered as a tool for monitoring progression in PSP, as the rate and region of atrophy may be heterogeneous within the disease.

MRI derived brain atrophy rates

- The results presented in chapter 9 of this thesis confirm that applying the brain boundary shift integral (BBSI) to assess brain volume change over time in PSP is a valid method of determining whole brain atrophy rates, and shows that PSP and MSA-P have similar rates of whole brain atrophy (1.2% $year^{-1}$ and 1.0% $year^{-1}$), approximately 2-3 times that seen in PD (0.6% year-1) or healthy age matched controls (0.4% $year^{-1}$).

However, according to power calculations, an impractical number of patients would be required in order to power trials of a disease modifying drug with whole brain atrophy rates as a marker of disease progression.

Regional brain atrophy rates as markers of progression

Manual segmentation techniques outlined in this study demonstrate that PSP and MSA-P have patterns of atrophy that distinguish them from PD, healthy controls and from each other.

- Comparing manually segmented regional volumes at two time points confirms that regional atrophy rates are disproportionate to whole brain atrophy.

- Applying the boundary shift integral (BSI) to these local regions (chapter 9) reduces measurement error and provides a more accurate operator dependent and less time-consuming method of calculating regional atrophy rates.
- Assessing the utility and feasibility of these measurements as markers of disease progression with a power calculation confirms that regional rather than whole brain atrophy rates, measured using the BSI, are more practical markers of disease progression as they require fewer study numbers (recruited patients) to generate statistical power in a trial of a putative disease modifying drug that slows atrophy rates.
- The shorter the scan interval, the greater the variance in measurements and this is likely to increase the number of subjects required to power a study. Increasing the interval to two years would reduce the variance considerably, possibly without a prohibitive increase in patient drop out.

Do regional brain volumes and atrophy rates have a clinical correlate?

The results outlined in Chapters 11 and 12 of this thesis identify associations between midbrain volume and clinical motor deficit, in keeping with the results of prior studies (Groschel et al). Additionally, SCP volume in the PSP patients was associated with motor deficit scored on the Hoehn and Yahr scale and with falls on the UPDRS II sub-section. Frontal volume was also related to neuropsychology test scores thought to be specific to that brain region, for example.

Far fewer associations were detected in the MSA-P patients but importantly; motor deficit was associated with pontine rather than midbrain volume.

- Progression in motor deficit is associated with a greater rate of midbrain atrophy in PSP and ponto-cerebellar atrophy in MSA-P.

- There were also associations between greater rates of frontal atrophy in PSP and worsening performance on neuropsychological tests of frontal cognitive function, although change in the FAB score was more strongly associated with midbrain atrophy rates, suggesting that progression in the severity of cognitive impairment occurs as a consequence of frontal cortical atrophy as well as subcortical pathology.

The associations described in Chapters 11 and 12 are likely to be limited by the number of patients studied, particularly in the MSA-P group. A larger prospective study of these patients rather than PSP patients would help to clarify the clinico-radiological associations of MSA-P.

Other points

Several important conclusions can be drawn from the preliminary studies outlined in chapters 3 and 5 in this thesis.

- PSP appears to be a heterogeneous disease. The analysis of clinical information collected from post mortem cases of PSP in a regional brain bank suggests that at least two distinct clinical phenotypes exist. The classical presentation with falls and a supranuclear gaze palsy, or Richardson's syndrome, and the more Parkinsonian variant, known as PSP-P.

It is important that these phenotypes have been identified: firstly as it means that cases of PSP previously labelled as atypical Parkinsonism (because they failed to fulfil the diagnostic criteria), may be identified and the true prevalence of PSP revised; and secondly because the two phenotypes have differing rates of progression and prognosis and may have differences in the underlying molecular pathological disease

mechanisms influencing the rate of progression. These factors will have implications in the interpretation of trials of disease modifying drugs. If MRI is used as a marker of disease progression, the increased variability in the measured atrophy rates would increase the numbers required to power a clinical trial or even mean that a real disease modifying effect is overlooked, if PSP patients are not adequately classified with appropriate matching of treatment and placebo arms.

While the results of fluid-SPM studies in this thesis suggest that in PSP there are differences in atrophy patterns between PSP patients with a severe and a less severe gaze palsy and with differing disease durations, further imaging research would be required to truly determine whether MRI derived rates of whole brain and regional atrophy differ between the two PSP phenotypes.

- The results of analysis of the FAB score in PSP, MSA-P and PD suggest that in addition to MRI features that support the clinical diagnosis of PSP, the clinician can apply a simple and repeatable bedside test of cognitive function that has been validated and quantified in PSP as well as other neurodegenerative diseases.

As such, the FAB correlates well with frontal tests of cognitive domains known to be affected in PSP, yet can be completed in a few minutes, providing a quantitative marker of the frontal cognitive impairment in an individual with PSP, helping to discriminate the more severe cognitive deficit of PSP from that of MSA-P and PD.

Comparison of the results of semi-quantitative pathological analysis with the results of fluid registered serial MRI in a single case of PSP (chapter 3.5.2) suggests that the

atrophy pattern detected during life using this method of MRI analysis has a clear pathological basis.

- Fluid registered MRI demonstrates a greater rate of atrophy in the motor cortex compared to the anterior frontal cortex with subsequent detection of more numerous tau positive structures, including glial inclusions and neuropil threads as well as neuronal loss. This preliminary study helps to confirm that regional atrophy detected using MRI arises as a consequence of direct pathological involvement of that region.

Finally, the results of the imaging studies suggest that PSP, MSA-P and PD patients with moderate to severe disease can tolerate serial MRI scans. The entire scan acquisition time for this study was under 20 minutes.

- A combination of careful explanation of the MRI procedure, along with experienced and helpful staff ensuring patient comfort, meant that over 80% of patients with what are considerably debilitating neurological disorders, profoundly affecting mobility, were able to complete the MRI protocol.

Factors which seemed to help maintain a high retention rate included: rapport with the patient and their carer, along with an explanation regarding the purpose of the study; an efficient means of transport to the scan appointment; performing all scans on the same scanner in a department with staff who have experience of patients with neurodegenerative disease and; using a scan protocol that is as short as possible. These simple factors, doubtless help to maintain minimal drop out in any clinical study using MRI.

Overall Conclusions

- Considering regional atrophy patterns together rather than independently on cross-sectional MRI may help to discriminate PSP more accurately during life.
- Marked atrophy of both the midbrain and the superior cerebellar peduncle are characteristic MRI features of PSP.
- The superior cerebellar peduncle should be visualised in axial and coronal planes and the midbrain assessed in the mid-sagittal plane. Considered together, these two regions may best identify PSP using MRI scans.
- Serial imaging showing excess atrophy rates above that expected for age may prove to be a useful means of detecting early neurodegeneration and patterns of regional atrophy which may help to differentiate degenerative diseases that present in a similar fashion.
- Regional atrophy rates derived from serial MRI scans rather than whole brain atrophy rates better discriminate PSP and MSA-P from PD, healthy controls and each other.
- These regional atrophy rates allow disease progression, as defined by tissue loss, to be accurately measured.
- Regional atrophy rates have clinically meaningful correlates and as such provide a feasible means of monitoring the effects of a putative disease-modifying drug.
- Definitive validation of MRI derived atrophy rates as a surrogate marker of progression can only occur when a confirmed disease modifying agent becomes available.

Publications

1. Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. *Williams DR, de Silva R,* ***Paviour DC****, Pittman A, Watt HC, Kilford L, Holton JL, Revesz T, Lees AJ.* Brain 2005 128(6): 1247-1258.

The clinical notes review of the PSP patients and data collection was conducted by Dr Paviour. Dr Williams reviewed these data and applied the statistical tests advised by Hilary Watt. Pathological studies and Tau Haplotype analysis were conducted by Drs Williams and de Silva, Alan Pittman, Dr Holton and Prof. Revesz.

2. Can the frontal assessment battery (FAB) differentiate bradykinetic rigid syndromes? Relation of the FAB to formal neuropsychological testing. ***D.C. Paviour, D.*** *Winterburn, S. Simmonds , G. Burgess, L. Wilkinson, N.C. Fox, A.J. Lees, M. Jahanshahi.* Neurocase, Volume 11, Number 4, August 2005: 274 – 282.

Dr Paviour recruited the patients, and conducted data analyses. David Winterburn, Saffron Simmonds, Gerald Burgess, Leonora Wilkinson and Marjan Jahanshahi conducted the psychological testing.

3. Pathological substrate for regional distribution of increased atrophy rates in progressive supranuclear palsy. ***D C Paviour****, J M Schott, J M Stevens, T Revesz, J L Holton, M N Rossor, A J Lees, N C Fox.* J Neurol. Neurosurg. Psychiatry. 2004 Dec; 75(12): 1772-5.

Dr Paviour drafted the paper with input from Dr Schott. Professor Revesz and Dr Holton conducted neuropathological analysis and Dr Stevens advised regarding radiological correlations. Professors Rossor, Lees and Fox helped to revise the paper.

4. Delineating the sites and progression of in vivo atrophy in multiple system atrophy using fluid-registered MRI. *Jonathan M. Schott, Jessica E. Simon, Nick C.*

Fox, Andrew P. King, M. Nadeem Khan, Lisa Cipolotti, ***Dominic C. Paviour****, John M. Stevens, Martin N. Rossor.* Mov Disord. 2003 Aug; 18(8): 955-8.

Dr Schott drafted the paper with input from Dr Paviour. The patient was seen during life by Dr Khan and Dr Simon. Dr Stevens provided image interpretation advise and Professor Rossor revised the manuscript.

5. Quantitative MRI measurement of superior cerebellar peduncle in progressive supranuclear palsy. ***D. C. Paviour****, S. L. Price, J. M. Stevens, A. J. Lees, N. C. Fox.* Neurology. 2005 Feb 22; 64(4): 675-9.

Dr Paviour recruited the patients, and performed image segmentation with advice from Shona Price. Dr Paviour drafted the manuscript. Dr Stevens reviewed and qualitatively assessed the imaging. Professors Lees and Fox revised the manuscript.

6. Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and Parkinson's disease. *Shona Price,* ***Dominic Paviour****, Rachael Scahill, John Stevens, Martin Rossor, Andrew Lees, Nick Fox.* Neuroimage. 2004 Oct; 23(2): 663-9.

Dr Paviour recruited the patients and helped to draft the manuscript. Image post processing and SPM analysis was performed by Shona Price who with Dr Stevens and Professors Rossor, Fox and Lees revised the manuscript.

7. Regional Brain Volumes Distinguish PSP, MSA-P and PD: MRI based Clinico-radiological Correlations. ***Paviour, D.C.*** *; Price, S.L.; Jahanshahi, M ; Lees, A.J.; Fox, N.C.* Movement Disorders 2006 (EPub ahead of print).

Dr Paviour recruited the patients and performed the clinical assessments as well as segmenting the regions of interest on the MRI scans. The clincoradiological associations were assessed by Dr Paviour. Other image analysis and post processing

was undertaken by Shona Price. Professors Jahanshahi, Lees and Fox helped to draft and revise the manuscript.

8. Longitudinal MRI in Progressive Supranuclear Palsy and Multiple System Atrophy: Rates and Regions of Atrophy. ***D.C. Paviour****; S.L. Price; M. Jahanshahi; A.J. Lees; N.C. Fox.* Brain 2006 (EPub ahead of print).

Dr Paviour recruited the patients and performed the clinical assessments as well as segmenting the regions of interest on the MRI scans. The clincoradiological associations were assessed by Dr Paviour. Other image analysis and post processing was undertaken by Shona Price. Professors Jahanshahi, Lees and Fox helped to draft and revise the manuscript.

Acknowledgements

I am grateful to the patients and carers who generously gave their time, allowing these research studies to be conducted. Without them, the project would have been impossible.

I would also like to acknowledge the help of the staff at the Institute of Neurology, involved in MRI acquisition, particularly Dave MacManus, for their expertise. Professor Marjan Jahanshahi and her research assistants were very helpful in designing and applying the neuropsychological test profile. Chris Frost provided expert statistical advice.

A particular acknowledgement goes to Shona Price whose expertise and patience regarding image post processing, particularly registration, segmentation and analysis was essential. Shona also helped along with Dr Rachael Scahill with input into the VBM and fluid-VBM studies.

Thanks go to close colleagues Andrew Evans, Dave Williams, Laura Moriyama, John Schott, Alison Godbolt, Basil Ridha, Hilary Archer and many others who have supported me throughout.

Particular thanks also to Julia and to my parents who have supported me during the completion of this thesis.

I am extremely indebted to Professor Andrew Lees and Professor Nick Fox who as my supervisors have provided positive and constructive criticism, guiding my research from its early beginnings.

The PSP Association, particularly Brigadier Koe who with his wife Sara founded the charity, deserve thanks for the funding of this research. Without Brigadier Koe's considerable drive and commitment to increasing the awareness of PSP, and his efforts towards fundraising, this research would not have been possible.

Appendix

Patient information sheet

A prospective clinicoradiological study in Progressive Supranuclear Palsy (PSP) using serial MRI scanning with registration.

(MRI scanning of the brain, over time in PSP)

PATIENT INFORMATION SHEET.

You are being invited to take part in a research study. Before you decide it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully, before you decide whether to take part.

What is the purpose of the study?

1. To help further understand the natural history of PSP.
2. To help develop a non invasive and safe diagnostic aid in the form of MRI scanning.
3. To help develop MRI as an aid to quantifying the results of therapeutic trials in PSP.

Why have I been chosen?

We have asked you to take part as you may have either PSP, multi system atrophy (MSA), Corticobasal degeneration (CBD) or Parkinson's disease (PD).
PSP is often misdiagnosed in its early stages or if the onset of the disease pattern is atypical. For this reason we are including in the study patients with conditions that PSP is sometimes mistaken for.
We hope to include at least 25 patients with PSP and 25 patients with either MSA or PD.

Do I have to take part?

It is up to you whether you decide to or not. If you decide to take part, you will be asked to sign a consent form. You would be free to withdraw from the study at any point should you so choose. If you decide not to take part or to withdraw at a future date, the care you receive will not be affected in any way.

What happens if I agree to take part?

You will be involved in the study for one year.
You will have a clinical examination by a neurologist at the time of entry into the study.
At this time, an MRI scan of your brain will be arranged and undertaken.
Magnetic Resonance Imaging (MRI) is a painless and safe technique, which can obtain detailed pictures of the brain. As the name implies, it uses magnetic fields to generate the pictures and unlike X-ray techniques there is no ionising radiation used. As long as people with any magnetic metal implants are excluded (such as pacemakers) there are no known risks.
The clinical assessment will include a detailed neuropsychological examination as well as a physical examination. The visit will take up most of the day including breaks for lunch.
Approximately one year later you will be seen in clinic again and a second MRI scan undertaken. We are happy to refund your travel expenses for these visits.
We will not be prescribing any medications as part of this trial.
You will be asked to fill out a questionnaire prior to scanning as cardiac pacemakers, containing magnetic metal, certain types of implants and artificial joints (not all types) will mean that MRI scanning cannot be undertaken.
We will inform you of the results of your scan should you so desire and if any unexpected abnormalities are found we would wish to notify you and your GP.

Are there any risks?
We do not envisage any risks to you in taking part in this study.
In the unlikely event of an unexpected, significant abnormality being found on the MRI brain scan, we would desire permission to notify yourself and your GP.

Are there any benefits?
Being involved in the study will not specifically help you as an individual. The results will hopefully play a part in helping diagnose and treat patients with PSP in the future.
Any information collected in this study will be kept strictly confidential
The information collected will be analysed over a period of approximately a year, leading to publication of the results. You will not be identified in any of the published results.

Who is funding the research?
The PSP association (Europe) is funding the costs of scanning and of the salary for one research assistant.

What if something goes wrong?
In the unlikely event of there being a problem with your participation and you wished to complain, the normal National Health Service (NHS) complaints mechanisms are available to you.

Other Points
The information obtained (clinical details regarding your illness, date of birth, name and address and the imaging data) will be stored on a computer and the study protocol has been reviewed by the National Hospital for Neurology and Neurosurgery and the Institute of Neurology Joint research Ethics Committee.

If you would like to take part then please sign and date the enclosed consent form and return it in the envelope provided. Alternatively contact Dr D. Paviour at the number or email address below.
Once entered into the study, a date for assessment and scanning will we allocated. If you desire, we will reimburse your travel costs.

Contact information Mon-Fri 9am-5pm
Dr Dominic Paviour

Other contacts
PSP Association (Europe) The Old Rectory Wappenham

www.pspeur.org Tel Fax

Version 1.3 - 12th May 2004

INFORMATION ON BRAIN DONATION

We appreciate your interest in becoming a brain donor. Although becoming a donor will not benefit you directly, it will contribute to the advancement of medical knowledge of brain disorders.

Please take time to read the following information carefully and discuss it with your family, friends or general practitioner if you wish.

An introduction to the Queen Square Brain Bank for Neurological Disorders

The Institute of Neurology and the National Hospital for Neurology and Neurosurgery have an international reputation as a centre of excellence for research into and treatment of neurological disorders. The Queen Square Brain Bank for Neurological Disorders (QSBB) has established a unique brain collection, which is used to study the effects of disease and to support research into disorders such as Parkinson's disease (PD), progressive supranuclear palsy (PSP), multiple system atrophy (MSA), dementia and dystonia. It has made a significant contribution to our understanding of the causes of neurological disease and in developing new treatments.

A brain donation is a gift. Because it is a very complicated organ, the whole brain is needed for diagnosis and research. In special cases we also examine the spinal cord, if it is thought to contribute to the disease process. Some tissue samples are preserved in fixative for diagnosis. Other tissue is frozen and kept at a low temperature so that it can be used for research over long periods of time. The research usually begins by examining the tissue under a microscope to identify any disease processes that have affected the brain. We then use specialist techniques to investigate possible abnormalities in brain and nerve cell proteins and in blood vessels in the brain. We may also analyse the DNA (genetic material) to look for abnormalities that may be specific to some diseases.

Brains are stored in a locked facility where only certain staff have access. To ensure all information remains confidential, the tissue is given a unique code number and records are held securely. To ensure anonymity the same unique code numbers are also used on photographs that may be taken of tissue specimens and microscopic slides.

Nervous system tissue is supplied to researchers in other universities both in the UK and world-wide. We also carry out some research with the support of industry. This is in line with the Institute of Neurology's policy of encouraging research that will lead to new therapies. In these instances intellectual property agreements are set up which means neither the scientists nor the donor's families gain a financial advantage.

The importance of correct diagnosis

Further information about many neurological disorders can only be gained by a detailed post-mortem and microscopic examination of the brain. For research to be precise a diagnosis made in life must be confirmed by neuropathological evaluation.

Results of research, which rely on collections of data from a large number of cases, are published in scientific journals and are not available on an individual basis. The anonymity of individual cases will be preserved in the research procedure and subsequent publications.

We would be happy to provide further information on the diagnosis and any other information available after the post-mortem, and this would be provided through the consultant physician in charge or the GP upon request by the next of kin. Relatives of the donor often gain comfort in knowing the nature of their loved one's illness. With development of scientific knowledge and improvements in neuropathological diagnosis further significant information relating to individual cases may become available in the future. At the request of the next of kin such information will be provided either by trained QSBB staff or appropriate medical or nursing personnel.

Brain donation by people with no neurological disease is of equal importance, providing essential 'control' tissue for comparison. Without this important gift research cannot go ahead.

How to become a brain donor

You should be over 18 years of age and a resident in the UK. You should discuss with your family your wish to donate your brain and if you wish, your spinal cord, to the QSBB following your death. Please complete and sign the enclosed "Declaration of Intent to Donate" form witnessed by your next of kin and send it to the Queen Square Brain Bank, Institute of Neurology, . If you are unable to complete this form yourself it is acceptable for a relative or appointed representative to do this for you.

Once your declaration form has been received by the QSBB you will be registered as a donor. We will send you a letter of welcome and a donor card, which you should carry at all times. You will also receive a self-assessment form, requesting information on your health and lifestyle. Potential donors with a neurological condition will be asked to complete this form on a yearly basis. This information will be stored on our computerised database so that essential clinical information is available to the pathologist.

The identity and the data collected about all registered donors remain strictly confidential.

At the time of your death your next of kin or appointed representative will be required to sign a consent form giving permission for the brain donation to take place.

Actions to be taken following the death of a donor

Following the death of a donor the QSBB must be contacted as soon as possible. Once informed of a donor's death by the next of kin, the QSBB will arrange for the donor's body to be transported to the nearest hospital where the removal of the brain and, if agreed, the spinal cord, can be performed. Before a post-mortem examination can be performed a QSBB consent form must be signed. The hospital at which the post-mortem examination is performed may require an additional local post-mortem consent form to be signed. Only the brain and the spinal cord are transferred to the QSBB. The donor's body can then be returned to the family's chosen funeral director and the usual funeral arrangements can then proceed. As brain donation does not require removal of the eyes or lead to disfigurement of the face, it does not prevent open casket or other traditional funeral arrangements. Removal of the spinal cord results in a long scar on either the front or back of the body, but this is not visible once the body has been prepared for the funeral. However, in view of this some donors may decide not to donate the spinal cord. All costs directly related to the removal of the brain will be met by the QSBB. Funeral costs must be met by the family.

For legal reasons some deaths require a post-mortem examination to be performed on behalf of HM Coroner. In these circumstances it is often still possible to arrange for the transfer and examination of the brain at the QSBB. This would require discussion between HM Coroner, QSBB medical staff and the next of kin at the time of death.

A letter to the next of kin will acknowledge receipt of the donor's brain.

Important notes

Even if registration was not completed during life, next-of-kin or appointed representative can still arrange for brain donation following death (in the knowledge that the donor was or would had been in agreement).

If any of your details change, for example your address, your general practitioner, diagnosis, etc. please notify the QSBB so that our records can be amended.

As time passes it may be necessary for you to alert new people to the bequest or remind others who may have forgotten.

Signing the declaration of intent form does not commit you to joining the scheme. If you change your mind in the future, you can withdraw your permission at any time.

DECLARATION OF INTENT TO DONATE MY BRAIN TO THE QUEEN SQUARE BRAIN BANK FOR NEUROLOGICAL DISORDERS

Version 1.2 – 1st September 2003

Please initial appropriate boxes

		YES	NO
1)	I have read the Information Sheet, dated, *12th May 2004* and have been given a copy to keep. I have had the opportunity to ask questions about brain donation and understand why the research is being undertaken.	☐	☐
2)	I agree to donate my whole brain for confirmation of clinical diagnosis and for research. I understand how the brain will be donated and the tissue retained. I understand that giving my brain for this research is voluntary and that I am free to withdraw my consent at any time without giving a reason.	☐	☐
3)	I agree to donate my spinal cord, if it is necessary, to aid confirmation of clinical diagnosis and for research purposes.	☐	☐
4)	I agree that the brain will be retained and used for:		
	a) medical research at the Institute of Neurology	☐	☐
	b) medical and scientific education	☐	☐
	c) the development of diagnostic tests and drugs with commercial organisations	☐	☐
	d) research with collaborators at other academic departments in the UK and worldwide.	☐	☐
	e) genetic (DNA) research purposes	☐	☐

5) I agree that members of the research team may look at my medical notes *(All information will remain confidential)* ☐ ☐

YES NO

6) I agree that my details can be stored on a database in accordance with the Data Protection Act 1998. ☐ ☐

YES **NO**

7) I understand that the results of genetic and other research will not be available on an individual basis. If and when such results are published they will be anonymised. ☐ ☐

YES **NO**

8) I agree that all decisions regarding the future use of my donated tissue will be made by members of staff of the Institute of Neurology, who will organise the lawful and respectful disposal of any remaining tissue after research studies are completed, in accordance with the guidelines of the Royal College of Pathologists. ☐ ☐

.....	.....	.....
Your Name (CAPITALS)	Your Signature	Date
.....	.....	.....
Witness's Name	Witness's Signature	Date

If you are unable to complete this form yourself, it is acceptable for a relative or appointed representative to do this for you.

Please complete this form and return in the enclosed envelope

Mini mental state examination

ORIENTATION

Maximum Score
Score
5 ( ) What is the (year)(season)(date)(day)(month)?
5 ( ) Where are we: (state)(county)(town)(hospital)(floor)?

REGISTRATION

3 ( ) Name 3 objects: 1 second to say each. Then ask the patient all 3 after you have said them. Give 1 point for each correct answer. Then repeat them until he learns all 3. Count trials and record.

Trials _____

ATTENTION AND CALCULATION

5 ( ) Serial 7's, 1 point for each correct. Stop after 5 answers. Alternatively, spell ÒworldÓ backwards.

RECALL

3 ( ) Ask for the 3 objects repeated above. One point for each correct answer

LANGUAGE

9 ( ) Identify and name a pencil and a watch (2 points)
Repeat the following "No ifs, ands, or buts" (1 point)
Follow a 3-stage command.
"Take a paper in your right hand, fold it in half, and put it on the floor" (3 points)
Read and obey the following:

Close your eyes (1 point)

Write a sentence (1 point)
Copy design (angles must intersect properly as shown)(1 point)

Total Score _____

INSTRUCTIONS FOR ADMINISTRATION

ORIENTATION

(1) Ask for the date. Then ask specifically for parts omitted, e.g., "can you tell me what season it is?" One point for each correct.

(2) Ask in turn "Can you tell me the name of this hospital?" (town, county, etc.) One point for each correct.

REGISTRATION

Ask the patient if you may test his memory. Then say the names of 3 unrelated objects, clearly and slowly, about 1 second for each. After you have said all 3, ask him to repeat them. This first repetition determines his score (0–3) but keep saying them until he can repeat all 3, up to 6 trials. If he does not eventually learn all 3, recall cannot be meaningfully tested.

ATTENTION AND CALCULATION

Ask the patient to begin with 100 and count backwards by 7. Stop after 5 subtractions (93, 86, 79, 72, 65). Score the total number of correct answers.

If the patient cannot or will not perform this task, ask him to spell the word "world" backwards. The score is the number of letters in correct order. E.g., dlrow = 5, dlorw = 3

RECALL

Ask the patient if he can recall the 3 words you previously asked him to remember. Score 0–3.

LANGUAGE

Naming: Show the patient a wrist watch and ask him what it is. Repeat for pencil. Score 0–2.

Repetition: Ask the patient to repeat the sentence after you. Allow only one trial. Score 0 or 1.

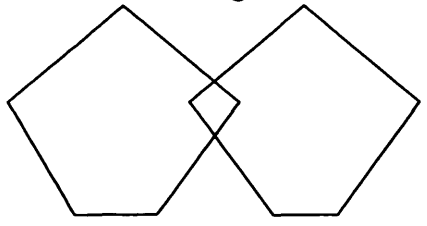
Three-stage command: Give the patient a piece of plain blank paper and repeat the command. Score 1 point for each part correctly executed.

Reading: On a blank piece of paper, print the sentence "close your eyes" in letters large enough for the patient to see clearly. Ask him to read it and do what it says. Score 1 point only if he actually closes his eyes.

Writing: Give the patient a blank piece of paper and ask him to write a sentence for you. Do not dictate a sentence; it is to be written spontaneously. It must contain a subject and verb and be sensible. Correct grammar and punctuation are not necessary.

Copying: On a clean piece of paper, draw intersecting pentagons, each side about 1 inch and ask him to copy it exactly as it is. All 10 angles must be present and 2 must intersect to score 1 point. Tremor and rotation are ignored.

Estimate the patient's level of sensorium along a continuum, from alert on the left to coma on the right.



CLOSE YOUR EYES

(Folstein *et al.*, 1975)

Unified Parkinson's Disease Rating Scale (UPDRS)

I. Mentation, Behavior and Mood

1. Intellectual Impairment:

   0—None.
   1—Mild. Consistent forgetfulness with partial recollection of events and no other difficulties.
   2—Moderate memory loss, with disorientation and moderate difficulty handling complex problems. Mild but definite impairment of function at home with need of occasional prompting.
   3—Severe memory loss with disorientation for time and often to place. Severe impairment in handling problems.
   4—Severe memory loss with orientation preserved to person only. Unable to make judgments or solve problems. Requires much help with personal care. Cannot be left alone at all.

2. Thought Disorder: (Due to dementia or drug intoxication)

   0—None.
   1—Vivid dreaming.
   2—"Benign" hallucinations with insight retained.
   3—Occasional to frequent hallucinations or delusions; without insight; could interfere with daily activities.
   4—Persistent hallucinations, delusions, or florid psychosis. Not able to care for self.

3. Depression:

   0—Not present.
   1—Periods of sadness or guilt greater than normal, never sustained for day or weeks.
   2—Sustained depression (1 week or more).
   3—Sustained depression with vegetative symptoms (insomnia, anorexia, weight loss, loss of interest).
   4—Sustained depression with vegetative symptoms and suicidal thoughts or intent.

4. Motivation/Initiative:

   0—Normal.
   1—Less assertive than usual; more passive.
   2—Loss of initiative or disinterest in elective (non-routine) activities.
   3—Loss of initiative or disinterest in day to day (routine) activities.
   4—Withdrawn, complete loss of motivation.

II. Activities of Daily Living (Determine for "on/off")

5. Speech:

   0—Normal.
   1—Mildly affected. No difficulty being understood.
   2—Moderately affected. Sometimes asked to repeat statements.
   3—Severely affected. Frequently asked to repeat statements.
   4—Unintelligible most of the time.

6. Salivation:

   0—Normal.
   1—Slight but definite excess of saliva in mouth; may have nighttime drooling.
   2—Moderately excessive saliva; may have minimal drooling.
   3—Marked excess of saliva with some drooling.
   4—Marked drooling, requires constant tissue or handkerchief.

7. Swallowing:

   0—Normal.
   1—Rare choking.

2—Occasional choking.
3—Requires soft food.
4—Requires NG tube or gastrotomy feeding.

8. Handwriting:

0—Normal.
1—Slightly slow or small.
2—Moderately slow or small; all words are legible.
3—Severely affected; not all words are legible.
4—The majority of words are not legible.

9. Cutting Food and Handling Utensils:

0—Normal.
1—Somewhat slow and clumsy, but no help needed.
2—Can cut most foods, although clumsy and slow; some help needed.
3—Food must be cut by someone, but can still feed slowly.
4—Needs to be fed.

10. Dressing:

0—Normal.
1—Somewhat slow, but no help needed.
2—Occasional assistance with buttoning, getting arms in sleeves.
3—Considerable help required, but can do some things alone.
4—Helpless.

11. Hygiene:

0—Normal.
1—Somewhat slow, but no help needed.
2—Needs help to shower or bathe, or very slow in hygienic care.
3—Requires assistance for washing, brushing teeth, combing hair, going to bathroom.
4—Foley catheter or other mechanical aids.

12. Turning in Bed and Adjusting Bed Clothes:

0—Normal.
1—Somewhat slow and clumsy, but no help needed.
2—Can turn alone or adjust sheets, but with great difficulty.
3—Can initiate, but not turn or adjust sheets alone.
4—Helpless.

13. Falling (unrelated to freezing):

0—None.
1—Rare falling.
2—Occasionally falls, less than once per day.
3—Falls an average of once daily.
4—Falls more than once daily.

14. Freezing When Walking:

    0—None.
    1—Rare freezing when walking; may have start-hesitation.
    2—Occasional freezing when walking.
    3—Frequent freezing. Occasionally falls from freezing.
    4—Frequent falls from freezing.

15. Walking:

    0—Normal.
    1—Mild difficulty. May not swing arms or may tend to drag leg.
    2—Moderate difficulty, but requires little or no assistance.
    3—Severe disturbance of walking, requiring assistance.
    4—Cannot walk at all, even with assistance.

16. Tremor:

    0—Absent.
    1—Slight and infrequently present.
    2—Moderate; bothersome to patient.
    3—Severe; interferes with many activities.
    4—Marked; interferes with most activities.

17. Sensory Complaints Related to Parkinsonism:

    0—None
    1—Occasionally has numbness, tingling, or mild aching.
    2—Frequently has numbness, tingling, or aching; not distressing.
    3—Frequent painful sensations.
    4—Excruciating pain.

III. Motor Examinations:

18. Speech:

    0—Normal.
    1—Slight loss of expression, diction, and/or volume.
    2—Monotone, slurred but understandable; moderately impaired.
    3—Marked impairment, difficult to understand.
    4—Unintelligible.

19. Facial Expression:

    0—Normal.
    1—Minimal hypomimia, could be normal "poker face."
    2—Slight but definitely abnormal diminution of facial expression.
    3—Moderate hypomimia; lips parted some of the time.
    4—Masked or fixed facies with severe of complete loss of facial expression; lips parted 1/4 inch or more.

20. Tremor at Rest: Head, L+R upper and lower limbs

    0—Absent.
    1—Slight and infrequently present.
    2—Mild in amplitude and persistent. Or moderate in amplitude, but only intermittently present.
    3—Moderate in amplitude and present most of the time.
    4—Marked in amplitude and present most of the time.

21. Action or Postural Tremor of Hands: Left and Right

0—Absent.

1—Slight; present with action.

2—Moderate in amplitude, present with action.

3—Moderate in amplitude with posture holding as well as action.

4—Marked in amplitude; interferes with feeding.

22. Rigidity: (Judged on passive movement of major joints with patient relaxed in sitting position. Cogwheeling to be ignored.) Neck, Left and Right upper and lower limbs

0—Absent.

1—Slight or detectable only when activated by mirror or other movements.

2—Mild to moderate.

3—Marked, but full range of motion easily achieved.

4—Severe, range of motion achieved with difficulty.

23. Finger Taps: (Patient taps thumb with index finger in rapid succession with widest amplitude possible, each hand separately.) Left and Right

0—Normal.

1—Mild slowing and/or reduction in amplitude.

2—Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.

3—Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement.

4—Can barely perform the task.

24. Hand Movements: (Patient opens and closes hand in rapid succession with widest amplitude possible, each hand separately.) Left and Right

0—Normal.

1—Mild slowing and/or reduction in amplitude.

2—Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.

3—Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement.

4—Can barely perform the task.

25. Rapid Alternating Movements of Hands: (Pronation-supination movements of hands, vertically or horizontally, with as large an amplitude as possible, both hands simultaneously.)

0—Normal.

1—Mild slowing and/or reduction in amplitude.

2—Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.

3—Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement.

4—Can barely perform the task.

26. Foot Agility: (Patient taps heel on ground in rapid succession, picking up entire foot. Amplitude should be about 3 inches) Left and Right.

0—Normal.

1—Mild slowing and/or reduction in amplitude.

2—Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.

3—Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement.

4—Can barely perform the task.

27. Arising from Chair: (Patient attempts to arise from a straight-back wood or metal chair with arms folded across chest.)

    0—Normal.

    1—Slow; or may need more than one attempt.

    2—Pushes self up from arms of seat.

    3—Tends to fall back and may have to try more than one time, but can get up without help.

    4—Unable to arise without help.

28. Posture:

    0—Normal erect.

    1—Not quite erect, slightly stooped posture; could be normal for older person.

    2—Moderately stooped posture, definitely abnormal; can be slightly leaning to one side.

    3—Severely stooped posture with kyphosis; can be moderately leaning to one side.

    4—Marked flexion with extreme abnormality of posture.

29. Gait:

    0—Normal

    1—Walks slowly, may shuffle with short steps, but no festination or propulsion.

    2—Walks with difficulty, but requires little or no assistance; may have some festination, short steps, or propulsion.

    3—Severe disturbance of gait, requiring assistance.

    4—Cannot walk at all, even with assistance.

30. Postural Stability: (Response to sudden posterior displacement produced by pull on shoulders while patient erect with eyes open and feet slightly apart. Patient is prepared.)

    0—Normal.

    1—Retropulsion, but recovers unaided.

    2—Absence of postural response; would fall if not caught by examiner.

    3—Very unstable, tends to lose balance spontaneously.

    4—Unable to stand without assistance.

31. Body Bradykinesia and Hypokinesia: (Combining slowness, hesitancy, decreased arm swing, small amplitude, and poverty of movement in general.)

    0—None.

    1—Minimal slowness, giving movement a deliberate character; could be normal for some persons. Possibly reduced amplitude.

    2—Mild degree of slowness and poverty of movement which is definitely abnormal. Alternatively, some reduced amplitude.

    3—Moderate slowness, poverty or small amplitude of movement.

    4—Marked slowness, poverty or small amplitude of movement.

IV. Complications of Therapy (in the past week)

A. Dyskinesias

32. Duration: What proportion of the waking day are dyskinesias present? (Historical information)

    0—None

    1—1–25% of day

    2—26–50% of day

    3—51–75% of day

    4—76–100% of day

33. Disability: How disabling are the dyskinesias? (Historical information; may be modified by office examination.)

    0—Not disabling

    1—Mildly disabling

    2—Moderately disabling

    3—Severely disabling

4—Completely disabling

34. Painful Dyskinesias: How painful are the dyskinesias?

    0—No painful dyskinesias
    1—Slight
    2—Moderate
    3—Severe
    4—Marked

35. Presence of Early Morning Dystonia: (Historical information)

    0—No
    1—Yes

B. Clinical Fluctuations

36. Are any "off" periods predictable as to timing after a dose of medication?

    0—No
    1—Yes

37. Are any "off" periods unpredictable as to timing after a dose of medication?

    0—No
    1—Yes

38. Do any of the "off" periods come on suddenly, e.g., over a few seconds?

    0—No
    1—Yes

39. What proportion of the waking day is patient "off" on average?

    0—None
    1—1–25% of day
    2—26–50% of day
    3—51–75% of day
    4—76–100% of day

C. Other Complications

40. Does the patient have anorexia, nausea, or vomiting?

    0—No
    1—Yes

41. Does the patient have any sleep disturbances, e.g., insomnia or hypersomnolence?

    0—No
    1—Yes

42. Does the patient have symptomatic orthostasis?

    0—No
    1—Yes

(Fahn *et al.*, 1987)

Hoehn and Yahr

Modified Hoehn and Yahr Staging

Stage 0—No signs of disease.
Stage 1—Unilateral disease.
Stage 1.5—Unilateral plus axial involvement.
Stage 2—Bilateral disease, without impairment of balance.
Stage 2.5—Mild bilateral disease with recovery on pull test.
Stage 3—Mild to moderate bilateral disease; some postural instability; physically independent.
Stage 4—Severe disability; still able to walk or stand unassisted.
Stage 5—Wheelchair bound or bedridden unless aided.

(Hoehn and Yahr, 1967)

Schwab and England ADL score

Modified Schwab and England Activities of Daily Living Scale

100%—Completely independent. Able to do all chores without slowness, difficulty, or impairment. Essentially normal. Unaware of any difficulty.

90%—Completely independent. Able to do all chores with some degree of slowness, difficulty, and impairment. Might take twice as long. Beginning to be aware of difficulty.

80%—Completely independent in most chores. Takes twice as long. Conscious of difficulty and slowness.

70%—Not completely independent. More difficulty with some chores. Three to four times as long in some. Must spend a large part of the day with chores.

60%—Some dependency. Can do most chores, but exceedingly slowly and with much effort. Errors; some impossible.

50%—More dependent. Help with half, slower, etc. Difficulty with everything.

40%—Very dependent. Can assist with all chores, but few alone.

30%—With effort, now and then does a few chores alone or begins alone. Much help needed.

20%—Nothing alone. Can be a slight help with some chores. Severe invalid.

10%—Total dependent, helpless. Complete invalid.

0%—Vegetative functions such as swallowing, bladder and bowel functions are not functioning. Bedridden.

Frontal assessment battery

1.Similarities (conceptualisation)

3 correct responses	☐	3	
2 correct responses	☐	2	
1 correct response	☐	1	
no correct response	☐	0	

2. Lexical fluency (mental flexibility)

more than 9 words	☐	3	
6-9 words	☐	2	
3-5 words	☐	1	
less than 3 words		☐	0

3. motor series (programming)

patient performs alone 6 correct consecutive series		☐	3
patient performs alone at least 3 consecutive series		☐	2
patient fails alone, but performs 3 correct consecutive series with the examiner		☐	1
patient cannot perform 3 correct consecutive series even with the examiner		☐	0

4. Conflicting instructions (sensitivity to interference)

no errors		☐	3
1 or 2 errors	☐	2	
more than 2 errors	☐	1	
the patient taps as the examiner at least 4 consecutive times		☐	0

5. Go-No go

no errors		☐	3
1 or 2 errors	☐	2	
more than 2 errors	☐	1	
the patient taps as the examiner at least 4 consecutive times		☐	0

6. Prehension behaviour (environmental autonomy)

the patient does not take the examiners hands		☐	3
the patient hesitates and asks what he/she has to do		☐	2
the patient takes the hands without hesitation		☐	1
the patient takes the examiners hands even after he or she has been told no to		☐	0
	total score		**/18**

1.Similarities

In what way are they alike:
A banana and an orange
A table and a chair
A tulip a rose and a daisy
Only categorical responses (fruits, furniture, flowers are considered correct)

2.Lexical fluency

"say as many words as you can beginning with the letter "S", no surnames or proper nouns"
Prompt the patient if no response in the first 5 seconds, e.g snake
If there are pauses then ask the question again
No repetitions, or variations allowed
Test duration is 60 seconds

3. Motor series

"look carefully at what I am doing" examiner performs "fist-edge-palm series"
"now you will do with your right hand, the same series, first with me then alone" examiner performs three times the series with the patient then tells the patient to continue alone.

4. Conflicting instructions

"tap twice when I tap once" to be sure that the patient has understood the instruction, a series of three trials is run 1-1-1.
"tap once when I tap twice" to be sure that the patient has understood the instruction, a series of three trials is run 2-2-2.
Then perform the following series: 1-1-2-1-2-2-2-1-1-2.

5. Go-no-go

"tap once when I tap once" to be sure that the patient has understood the instruction, a series of three trials is run 1-1-1.
"do not tap when I tap twice" to be sure that the patient has understood the instruction, a series of three trials is run 2-2-2.
Then perform the following series: 1-1-2-1-2-2-2-1-1-2.

6. Prehension behaviour

The examiner is seated in front of the patient. Place the patient's hands palms up on his/her knees. Without saying or looking at the patient, the examiner brings his/her hands close to the patient's hands and touches the patient's palms of both hands, to see if he/she will spontaneously take them. If the patient takes the hands, the examiner will try again after asking him/her: "now do not take my hands".

(Dubois *et al.*, 2000)

Gaze palsy score

Speed of voluntary upward saccades

Normal	0
Mild slowness	1
Definite slowness	2
Severe slowness	3

Amplitude of voluntary upward saccades

Not hypometric	0
Mild limitation	1
Moderate limitation	2
Severe limitation	3

Speed of voluntary downward saccades

Normal speed	0
Mild slowness	1
Definite slowness	2
Severe slowness	3

Amplitude of voluntary downward saccades

Not hypometric	0
Mild limitation	1
Moderate limitation	2
Severe limitation	3

Speed of voluntary left and right saccades

Normal speed	0
Mild slowness	1
Definite slowness	2
Severe slowness	3

Amplitude of voluntary left and right saccades

Not hypometric	0
Mild limitation	1
Moderate limitation	2
Severe limitation	3

Eyelid dysfunction

None	0
Mild inhibition of opening or closing or mild blepharospasm: no visual disability	1
Moderate lid-opening inhibition or blepharospasm causing partial visual disability	2
Functional blindness or near blindness because of involuntary eyelid closure	3

TOTAL = /21

References

Aarsland D, Andersen K, Larsen JP, Lolk A, Kragh-Sorensen P. Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. Arch Neurol 2003; 60: 387-392.

Aarsland D, Larsen JP, Lim NG, Janvin C, Karlsen K, Tandberg E *et al.* Range of neuropsychiatric disturbances in patients with Parkinson's disease. J Neurol Neurosurg Psychiatry 1999; 67: 492-496.

Aarsland D, Tandberg E, Larsen JP, Cummings JL. Frequency of dementia in Parkinson disease. Arch Neurol 1996; 53: 538-542.

Adachi M, Kawanami T, Ohshima H, Sugai Y, Hosoya T. Morning glory sign: a particular MR finding in progressive supranuclear palsy. Magn Reson Med Sci. 2004; 3: 125-132.

Aiba I, Hashizume Y, Yoshida M, Okuda S, Murakami N, Ujihira N. Relationship between brainstem MRI and pathological findings in progressive supranuclear palsy--study in autopsy cases. J Neurol Sci. 1997; 152: 210-217.

Albert ML, Feldman RG, Willis AL. The 'subcortical dementia' of progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 1974; 37: 121-130.

Alexander GE, Crutcher MD. Basal ganglia-thalamo-cortical circuits: Parallel substrates for motor, oculomotor, "prefrontal " and "limbic" functions. The Prefrontal

cortex:Its structure, function and pathology. Amsterdam: Elsevier Science Publishers; 1990: 119-43.

Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu.Rev.Neurosci 1986; 9: 357-381.

Angibaud G, Gaultier C, Rascol O. Atypical parkinsonism and Annonaceae consumption in New Caledonia. Mov Disord 2004; 19: 603-604.

Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML *et al.* The Frontal Assessment Battery (FAB): normative values in an Italian population sample. Neurol Sci. 2005; 26: 108-116.

Asanuma C, Thach WR, Jones EG. Anatomical evidence for segregated focal groupings of efferent cells and their terminal ramifications in the cerebellothalamic pathway of the monkey. Brain Res. 1983; 286: 267-297.

Asato R, Akiguchi I, Masunaga S, Hashimoto N. Magnetic resonance imaging distinguishes progressive supranuclear palsy from multiple system atrophy. J Neural Transm. 2000; 107: 1427-1436.

Ashburner J, Csernansky JG, Davatzikos C, Fox NC, Frisoni GB, Thompson PM. Computer-assisted imaging to assess brain structure in healthy and diseased brains. Lancet Neurol 2003; 2: 79-88.

Ashburner J, Friston KJ. Voxel-based morphometry--the methods. Neuroimage. 2000; 11: 805-821.

Averbuch-Heller L, Paulson GW, Daroff RB, Leigh RJ. Whipple's disease mimicking progressive supranuclear palsy: the diagnostic value of eye movement recording. J Neurol Neurosurg Psychiatry 1999; 66: 532-535.

Barbeau A. [Plurisystematic degeneration of the neuraxis. Steele-Richardson-Olszewski syndrome]. Union Med Can. 1965; 94: 715-718.

Barnes J, Scahill RI, Boyes RG, Frost C, Lewis EB, Rossor CL *et al.* Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates. NeuroImage 2004; 23: 574-581.

Beck AT, Steer RA. Beck anxiety inventory, manual. San Antonio: Psychological corporation; 1990.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen.Psychiatry 1961; 4: 561-571.

Benton AL. Differential behavioral effects in frontal lobe disease. Neuropsychologia 1968; 6: 53-60.

Bentson JR, Keesey JC. Pneumoencephalography of progressive supranuclear palsy. Radiology 1974; 113: 89-94.

Bergeron C, Pollanen MS, Weyer L, Lang AE. Cortical degeneration in progressive supranuclear palsy. A comparison with cortical-basal ganglionic degeneration. J Neuropathol.Exp.Neurol 1997; 56: 726-734.

Bhatia KP, Marsden CD. The behavioural and motor consequences of focal lesions of the basal ganglia in man. Brain 1994; 117 (Pt 4): 859-876.

Bhattacharya K, Saadia D, Eisenkraft B, Yahr M, Olanow W, Drayer B *et al.* Brain magnetic resonance imaging in multiple-system atrophy and Parkinson disease: a diagnostic algorithm. Arch.Neurol. 2002; 59: 835-842.

Bigio EH, Brown DF, White CL, III. Progressive supranuclear palsy with dementia: cortical pathology. J Neuropathol.Exp.Neurol 1999; 58: 359-364.

Birdi S, Rajput AH, Fenton M, Donat JR, Rozdilsky B, Robinson C *et al.* Progressive supranuclear palsy diagnosis and confounding features: report on 16 autopsied cases. Mov Disord 2002; 17: 1255-1264.

Blin J, Baron JC, Dubois B, Pillon B, Cambon H, Cambier J *et al.* Positron emission tomography study in progressive supranuclear palsy. Brain hypometabolic pattern and clinicometabolic correlations. Arch.Neurol 1990; 47: 747-752.

Braak H, Jellinger K, Braak E, Bohl J. Allocortical neurofibrillary changes in progressive supranuclear palsy. Acta Neuropathol.(Berl) 1992; 84: 478-483.

Brenneis C, Boesch SM, Egger KE, Seppi K, Scherfler C, Schocke M *et al.* Cortical atrophy in the cerebellar variant of multiple system atrophy: A voxel-based morphometry study. Mov Disord 2005.

Brenneis C, Seppi K, Schocke M, Benke T, Wenning GK, Poewe W. Voxel based morphometry reveals a distinct pattern of frontal atrophy in progressive supranuclear palsy. Journal of Neurology, Neurosurgery, and Psychiatry 2004; 75: 246-249.

Brenneis C, Seppi K, Schocke MF, Muller J, Luginger E, Bosch S *et al.* Voxel-based morphometry detects cortical atrophy in the Parkinson variant of multiple system atrophy. Mov Disord. 2003; 18: 1132-1138.

Brown RG, Marsden CD. 'Subcortical dementia': the neuropsychological evidence. Neuroscience 1988; 25: 363-387.

Brucke T, Asenbaum S, Pirker W, Djamshidian S, Wenger S, Wober C *et al.* Measurement of the dopaminergic degeneration in Parkinson's disease with [123I] beta-CIT and SPECT. Correlation with clinical findings and comparison with multiple system atrophy and progressive supranuclear palsy. J Neural Transm.Suppl 1997; 50: 9-24.

Burk K, Skalej M, Dichgans J. Pontine MRI hyperintensities ("the cross sign") are not pathognomonic for multiple system atrophy (MSA). Mov Disord 2001; 16: 535.

Burton EJ, McKeith IG, Burn DJ, Williams ED, Obrien JT. Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with Lewy bodies and controls. Brain 2004; 127: 791-800.

Burton EJ, Karas G, Paling SM, Barber R, Williams ED, Ballard CG *et al.* Patterns of cerebral atrophy in dementia with Lewy bodies using voxel- based morphometry. Neuroimage. 2002; 17: 618-630.

Campdelacreu J, Kumru H, Tolosa E, Valls-Sole J, Benabarre A. Progressive supranuclear palsy syndrome induced by clebopride. Mov Disord 2004; 19: 482-484.

Caparros-Lefebvre D, Sergeant N, Lees A, Camuzat A, Daniel S, Lannuzel A *et al.* Guadeloupean parkinsonism: a cluster of progressive supranuclear palsy- like tauopathy. Brain 2002; 125: 801-811.

Chan D, Fox NC, Jenkins R, Scahill RI, Crum WR, Rossor MN. Rates of global and regional cerebral atrophy in AD and frontotemporal dementia. Neurology 2001a; 57: 1756-1763.

Chan D, Fox NC, Scahill RI, Crum WR, Whitwell JL, Leschziner G *et al.* Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Ann.Neurol. 2001b; 49: 433-442.

Collins SJ, Ahlskog JE, Parisi JE, Maraganore DM. Progressive supranuclear palsy: neuropathologically based diagnostic clinical criteria. J Neurol Neurosurg Psychiatry 1995; 58: 167-173.

Conrad C, Andreadis A, Trojanowski JQ, Dickson DW, Kang D, Chen X *et al.* Genetic evidence for the involvement of tau in progressive supranuclear palsy. Ann.Neurol 1997; 41: 277-281.

Cordato NJ, Duggins AJ, Halliday GM, Morris JGL, Pantelis C. Clinical deficits correlate with regional cerebral atrophy in progressive supranuclear palsy. Brain 2005; 128: 1259-1266.

Cordato NJ, Pantelis C, Halliday GM, Velakoulis D, Wood SJ, Stuart GW *et al.* Frontal atrophy correlates with behavioural changes in progressive supranuclear palsy. Brain 2002; 125: 789-800.

Cordato NJ, Halliday GM, Harding AJ, Hely MA, Morris JG. Regional brain atrophy in progressive supranuclear palsy and Lewy body disease. Ann.Neurol 2000; 47: 718-728.

Crank J. The Mathematics of Diffusion. Oxford: Oxford Science Clarendon Press; 1998.

Crum WR, Scahill RI, Fox NC. Automated hippocampal segmentation by regional fluid registration of serial MRI: validation and application in Alzheimer's disease. Neuroimage. 2001; 13: 847-855.

Cubo E, Stebbins GT, Golbe LI, Nieves A, Leurgans S, Goetz CG *et al.* Application of the Unified Parkinson's Disease Rating Scale in progressive supranuclear palsy: factor analysis of the motor scale. Mov Disord. 2000; 15: 276-279.

Dalton CM, Brex PA, Jenkins R, Fox NC, Miszkiel KA, Crum WR *et al.* Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. Journal of Neurology, Neurosurgery, and Psychiatry 2002; 73: 141-147.

Daniel SE, de Bruin VM, Lees AJ. The clinical and pathological spectrum of Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy): a reappraisal. Brain 1995; 118 (Pt 3): 759-770.

Davis PH, Bergeron C, McLachlan DR. Atypical presentation of progressive supranuclear palsy. Ann.Neurol 1985; 17: 337-343.

de Bruin VM, Lees AJ, Daniel SE. Diffuse Lewy body disease presenting with supranuclear gaze palsy, parkinsonism, and dementia: a case report. Mov Disord. 1992; 7: 355-358.

Dickens C CW. The Lazy Tour of Two Idle Apprentices In: No thoroughfare and other stories. Stroud: Alan Sutton; 1990.

Drayer BP, Olanow W, Burger P, Johnson GA, Herfkens R, Riederer S. Parkinson plus syndrome: diagnosis using high field MR imaging of brain iron. Radiology 1986; 159: 493-498.

Dubas F, Gray F, Escourolle R. [Steele-Richardson-Olszewski disease without ophthalmoplegia. 6 clinico- anatomic cases]. Rev.Neurol (Paris) 1983; 139: 407-416.

Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. Neurology 2000; 55: 1621-1626.

Dubois B, Pillon B, Legault F, Agid Y, Lhermitte F. Slowing of cognitive processing in progressive supranuclear palsy. A comparison with Parkinson's disease. Arch.Neurol 1988; 45: 1194-1199.

Duguid JR, De La PR, DeGroot J. Magnetic resonance imaging of the midbrain in Parkinson's disease. Ann.Neurol. 1986; 20: 744-747.

Dum RP, Strick PL. An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. J Neurophysiol. 2003; 89: 634-639.

Duvoisin RC. Progressive supranuclear palsy: clinical and research approaches. In: Litvan I AY, editor. New York: Oxford University Press; 1992: 15-33.

Emre M. Dementia associated with Parkinson's disease. Lancet Neurol 2003; 2: 229-237.

Esmonde T, Giles E, Xuereb J, Hodges J. Progressive supranuclear palsy presenting with dynamic aphasia. J Neurol Neurosurg Psychiatry 1996; 60: 403-410.

Ezekiel F, Chao L, Kornak J, Du AT, Cardenas V, Truran D *et al.* Comparisons between global and focal brain atrophy rates in normal aging and Alzheimer disease: Boundary Shift Integral versus tracing of the entorhinal cortex and hippocampus. Alzheimer Dis Assoc Disord 2004; 18: 196-201.

Fahn S, Elton R, Members of the UPDRS Development Committee. Recent Developments in Parkinson's Disease, Vol 2. Florham Park, NJ.: Macmillan Health Care Information; 1987.

Fearnley JM, Revesz T, Brooks DJ, Frackowiak RS, Lees AJ. Diffuse Lewy body disease presenting with a supranuclear gaze palsy. J Neurol Neurosurg Psychiatry 1991; 54: 159-161.

Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern.Med 1996; 125: 605-613.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr.Res. 1975; 12: 189-198.

Fox NC, Schott JM. Imaging cerebral atrophy: normal ageing to Alzheimer's disease. Lancet 2004; 363: 392-394.

Fox NC, Cousens S, Scahill R, Harvey RJ, Rossor MN. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. Arch.Neurol. 2000; 57: 339-344.

Fox NC, Scahill RI, Crum WR, Rossor MN. Correlation between rates of brain atrophy and cognitive decline in AD. Neurology 1999; 52: 1687-1689.

Fox NC, Freeborough PA. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. J.Magn Reson.Imaging 1997; 7: 1069-1075.

Fox NC, Freeborough PA, Rossor MN. Visualisation and quantification of rates of atrophy in Alzheimer's disease. Lancet 1996; 348: 94-97.

Freeborough PA, Fox NC. Modeling brain deformations in Alzheimer disease by fluid registration of serial 3D MR images. J.Comput.Assist.Tomogr. 1998; 22: 838-843.

Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans.Med.Imaging 1997; 16: 623-629.

Freeborough PA, Fox NC, Kitney RI. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. Comput.Methods Programs Biomed. 1997; 53: 15-25.

Freeborough PA, Woods RP, Fox NC. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. J.Comput.Assist.Tomogr. 1996; 20: 1012-1022.

Friston KJ. Commentary and opinion: II. Statistical parametric mapping: ontology and current issues. J Cereb.Blood Flow Metab 1995; 15: 361-370.

Friston KJ, Holmes A, Poline JB, Price CJ, Frith CD. Detecting activations in PET and fMRI: levels of inference and power. NeuroImage 1996; 4: 223-235.

Gibb GM, de Silva R, Revesz T, Lees AJ, Anderton BH, Hanger DP. Differential involvement and heterogeneous phosphorylation of tau isoforms in progressive supranuclear palsy. Brain Res.Mol.Brain Res. 2004; 121: 95-101.

Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988; 51: 745-752.

Gilman S, Low PA, Quinn N, Albanese A, Ben Shlomo Y, Fowler CJ *et al.* Consensus statement on the diagnosis of multiple system atrophy. J Neurol Sci. 1999; 163: 94-98.

Gimenez-Roldan S, Mateo D, Benito C, Grandas F, Perez-Gilabert Y. Progressive supranuclear palsy and corticobasal ganglionic degeneration: differentiation by clinical features and neuroimaging techniques. J Neural Transm.Suppl 1994; 42: 79-90.

Golbe LI. Progressive supranuclear palsy. In; Parkinson's disease and movement disorders. In: Jankovic J TE, editor. Baltimore: Williams & Wilkins; 1993: 145-61.

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage. 2001; 14: 21-36.

Grafman J, Litvan I, Stark M. Neuropsychological features of progressive supranuclear palsy. Brain Cogn 1995; 28: 311-320.

Gronwall D, Wrightson P. Memory and information processing capacity after closed head injury. J Neurol Neurosurg Psychiatry 1981; 44: 889-895.

Groschel K, Hauser TK, Luft A, Patronas N, Dichgans J, Litvan I *et al.* Magnetic resonance imaging-based volumetry differentiates progressive supranuclear palsy from corticobasal degeneration. NeuroImage 2004; 21: 714-724.

Gunter JL, Shiung MM, Manduca A, Jack CR, Jr. Methodological considerations for measuring rates of brain atrophy. J Magn Reson.Imaging 2003; 18: 16-24.

Haldeman S, Goldman JW, Hyde J, Pribram HF. Progressive supranuclear palsy, computed tomography, and response to antiparkinsonian drugs. Neurology 1981; 31: 442-445.

Halliday GM, Macdonald V, Henderson JM. A comparison of degeneration in motor thalamus and cortex between progressive supranuclear palsy and Parkinson's disease. Brain 2005; awh596.

Hardman CD, Halliday GM, McRitchie DA, Morris JG. The subthalamic nucleus in Parkinson's disease and progressive supranuclear palsy. J Neuropathol.Exp.Neurol 1997; 56: 132-142.

Harris ML, Julyan P, Kulkarni B, Gow D, Hobson A, Hastings D *et al.* Mapping metabolic brain activation during human volitional swallowing: a positron emission tomography study using [lsqb]18F[rsqb]fluorodeoxyglucose. 2005; 25: 520-526.

Hauw JJ, Daniel SE, Dickson D, Horoupian DS, Jellinger K, Lantos PL *et al.* Preliminary NINDS neuropathologic criteria for Steele-Richardson- Olszewski syndrome (progressive supranuclear palsy). Neurology 1994; 44: 2015-2019.

Hayflick SJ, Westaway SK, Levinson B, Zhou B, Johnson MA, Ching KH *et al.* Genetic, clinical, and radiographic delineation of Hallervorden-Spatz syndrome. N.Engl.J Med 2003; 348: 33-40.

Henderson JM, Carpenter K, Cartwright H, Halliday GM. Loss of thalamic intralaminar nuclei in progressive supranuclear palsy and Parkinson's disease: clinical and therapeutic implications. Brain 2000; 123 (Pt 7): 1410-1421.

Hirano A, Kurland LT, Krooth RS, Lessell S. Parkinsonism-dementia complex, an endemic disease on the island of Guam. I. Clinical features. Brain 1961a; 84: 642-661.

Hirano A, Malamud N, Kurland LT. Parkinsonism-dementia complex, an endemic disease on the island of Guam. II. Pathological features. Brain 1961b; 84: 662-679.

Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967; 17: 427-442.

Holsapple JW, Preston JB, Strick PL. The origin of thalamic inputs to the "hand" representation in the primary motor cortex. J Neurosci 1991; 11: 2644-2654.

Hoover JE, Strick PL. The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1. J Neurosci 1999; 19: 1446-1463.

Horimoto Y, Aiba I, Yasuda T, Ohkawa Y, Katayama T, Yokokawa Y *et al.* Longitudinal MRI study of multiple system atrophy - when do the findings appear, and what is the course? J Neurol 2002; 249: 847-854.

Horimoto Y, Aiba I, Yasuda T, Ohkawa Y, Katayama T, Yokokawa Y *et al.* Cerebral atrophy in multiple system atrophy by MRI. J Neurol Sci. 2000; 173: 109-112.

Hosaka K, Ishii K, Sakamoto S, Mori T, Sasaki M, Hirono N *et al.* Voxel-based comparison of regional cerebral glucose metabolism between PSP and corticobasal degeneration. J Neurol Sci. 2002; 199: 67-71.

Hughes AJ, Daniel SE, Ben Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain 2002; 125: 861-870.

Hughes AJ, Daniel SE, Blankson S, Lees AJ. A clinicopathologic study of 100 cases of Parkinson's disease. Arch Neurol 1993; 50: 140-148.

Jack CR, Jr., Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS *et al.* Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 2004; 62: 591-600.

Jack CR, Jr., Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ *et al.* Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 2000; 55: 484-489.

Jack CR, Jr., Petersen RC, Xu YC, Waring SC, O'Brien PC, Tangalos EG *et al.* Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 1997; 49: 786-794.

Jenkins R, Fox NC, Rossor AM, Harvey RJ, Rossor MN. Intracranial volume and Alzheimer disease: evidence against the cerebral reserve hypothesis. Arch.Neurol. 2000; 57: 220-224.

Josephs KA, Ishizawa T, Tsuboi Y, Cookson N, Dickson DW. A clinicopathological study of vascular progressive supranuclear palsy: a multi-infarct disorder presenting as progressive supranuclear palsy. Arch.Neurol 2002; 59: 1597-1601.

Kanazawa M, Shimohata T, Terajima K, Onodera O, Tanaka K, Tsuji S *et al.* Quantitative evaluation of brainstem involvement in multiple system atrophy by diffusion-weighted MR imaging. J Neurol 2004; 251: 1121-1124.

Kantarci K, Petersen RC, Boeve BF, Knopman DS, Weigand SD, O'Brien PC *et al.* DWI predicts future progression to Alzheimer disease in amnestic mild cognitive impairment. Neurology 2005; 64: 902-904.

Kato N, Arai K, Hattori T. Study of the rostral midbrain atrophy in progressive supranuclear palsy. J Neurol Sci. 2003; 210: 57-60.

Kim SG, Ugurbil K, Strick PL. Activation of a cerebellar output nucleus during cognitive processing. Science 1994; 265: 949-951.

Kimura D, Barnett HJ, Burkhart G. The psychological test pattern in progressive supranuclear palsy. Neuropsychologia 1981; 19: 301-306.

Kirkwood BR, Sterne JAC. Calculation of required sample size. Medical Statistics. Malden, Massachusetts: Blackwell Science; 2003: 413-28.

Kluin K, Gilman S, Foster N, Sima A, D'Amato C, Bruch L *et al.* Neuropathological correlates of dysarthria in progressive supranuclear palsy. Arch.Neurol 2001; 58: 265-269.

Konagaya M, Konagaya Y, Sakai M, Matsuoka Y, Hashizume Y. Progressive cerebral atrophy in multiple system atrophy. J Neurol Sci. 2002; 195: 123-127.

Konagaya M, Sakai M, Matsuoka Y, Konagaya Y, Hashizume Y. Multiple system atrophy with remarkable frontal lobe atrophy. Acta Neuropathol (Berl) 1999a; 97: 423-428.

Konagaya M, Sakai M, Matsuoka Y, Konagaya Y, Hashizume Y. Multiple system atrophy with remarkable frontal lobe atrophy. Acta Neuropathol.(Berl) 1999b; 97: 423-428.

Kopelman MD. Frontal dysfunction and memory deficits in the alcoholic Korsakoff syndrome and Alzheimer-type dementia. Brain 1991; 114 (Pt 1A): 117-137.

Lange KW, Tucha O, Alders GL, Preier M, Csoti I, Merz B *et al.* Differentiation of parkinsonian syndromes according to differences in executive functions. J Neural Transm. 2003; 110: 983-995.

Larner AJ. Did Charles Dickens describe progressive supranuclear palsy in 1857? Mov Disord. 2002; 17: 832-833.

Lees AJ. The Steele Richardson Olszewski syndrome. In: Marsden CD, Fahn S, editors. Movement Disorders 2. Butterworths; 1987: 272-87.

Leigh RJ, Kennard C. Using saccades as a research tool in the clinical neurosciences. Brain 2004; 127: 460-477.

Leigh RJ, Riley DE. Eye movements in parkinsonism: it's saccadic speed that counts. Neurology 2000; 54: 1018-1019.

Leiguarda RC, Pramstaller PP, Merello M, Starkstein S, Lees AJ, Marsden CD. Apraxia in Parkinson's disease, progressive supranuclear palsy, multiple system atrophy and neuroleptic-induced parkinsonism. Brain 1997; 120 (Pt 1): 75-90.

Levin BE, Katzen HL. Early cognitive changes and nondementing behavioral abnormalities in Parkinson's disease. Adv.Neurol 1995; 65: 85-95.

Lewis EB, Fox NC. Correction of differential intensity inhomogeneity in longitudinal MR images. NeuroImage 2004; 23: 75-83.

Litvan I, Bhatia KP, Burn DJ, Goetz CG, Lang AE, McKeith I *et al.* SIC Task Force appraisal of clinical diagnostic criteria for parkinsonian disorders. Mov Disord. 2003; 18: 467-486.

Litvan I, Grimes DA, Lang AE, Jankovic J, McKee A, Verny M *et al.* Clinical features differentiating patients with postmortem confirmed progressive supranuclear palsy and corticobasal degeneration. J Neurol 1999; 246 Suppl 2: II1-II5.

Litvan I. The clinical and pathologic hallmarks of progressive supranuclear palsy. Curr.Opin.Neurol 1997; 10: 346-350.

Litvan I, Agid Y, Goetz C, Jankovic J, Wenning GK, Brandel JP *et al.* Accuracy of the clinical diagnosis of corticobasal degeneration: a clinicopathologic study. Neurology 1997; 48: 119-125.

Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC *et al.* Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. Neurology 1996a; 47: 1-9.

Litvan I, Hauw JJ, Bartko JJ, Lantos PL, Daniel SE, Horoupian DS *et al.* Validity and reliability of the preliminary NINDS neuropathologic criteria for progressive supranuclear palsy and related disorders. J Neuropathol.Exp.Neurol 1996b; 55: 97-105.

Litvan I, Mangone CA, McKee A, Verny M, Parsa A, Jellinger K *et al.* Natural history of progressive supranuclear palsy (Steele-Richardson- Olszewski syndrome) and clinical predictors of survival: a clinicopathological study. J Neurol Neurosurg Psychiatry 1996c; 60: 615-620.

Litvan I, Mega MS, Cummings JL, Fairbanks L. Neuropsychiatric aspects of progressive supranuclear palsy. Neurology 1996d; 47: 1184-1189.

Litvan I. Cognitive disturbances in progressive supranuclear palsy. J Neural Transm.Suppl 1994; 42: 69-78.

Litvan I, Grafman J, Gomez C, Chase TN. Memory impairment in patients with progressive supranuclear palsy. Arch.Neurol. 1989; 46: 765-767.

Lo CD, Lunetta C, Piccoli F, La B, V. How specific are the pontine MRI hyperintensities (the cross sign)? Eur J Neurol 2003; 10: 108-109.

Maher ER, Lees AJ. The clinical features and natural history of the Steele-Richardson- Olszewski syndrome (progressive supranuclear palsy). Neurology 1986; 36: 1005-1008.

Maher ER, Smith EM, Lees AJ. Cognitive deficits in the Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). J Neurol Neurosurg Psychiatry 1985; 48: 1234-1239.

Mamata H, Mamata Y, Westin CF, Shenton ME, Kikinis R, Jolesz FA *et al.* High-resolution line scan diffusion tensor MR imaging of white matter fiber tract anatomy. AJNR Am.J Neuroradiol. 2002; 23: 67-75.

Maraganore DM, Anderson DW, Bower JH, McDonnell SK, Rocca WA. Autopsy patterns for Parkinson's disease and related disorders in Olmsted County, Minnesota. Neurology 1999; 53: 1342-1344.

Marin RS. Apathy: Concept, Syndrome, Neural Mechanisms, and Treatment. Semin.Clin Neuropsychiatry 1996; 1: 304-314.

Marin RS. Differential diagnosis and classification of apathy. American Journal of Psychiatry 1990; 147: 22-30.

Marx JJ, Thoemke F, Mika-Gruettner A, Fitzek S, Vucurevic G, Urban PP *et al.* [Diffusion-weighted MRT in vertebrobasilar ischemia. Application, sensitivity, and prognostic value]. Nervenarzt 2004; 75: 341-346.

Masliah E, Hansen LA, Quijada S, DeTeresa R, Alford M, Kauss J *et al.* Late onset dementia with argyrophilic grains and subcortical tangles or atypical progressive supranuclear palsy? Ann.Neurol 1991; 29: 389-396.

Masucci EF, Borts FT, Perl SM, Wener L, Schwankhaus J, Kurtzke JF. MR vs CT in progressive supranuclear palsy. Comput.Med.Imaging Graph. 1995; 19: 361-368.

Masucci EF, Borts FT, Smirniotopoulos JG, Kurtzke JF, Schellinger D. Thin-section CT of midbrain abnormalities in progressive supranuclear palsy. AJNR Am J Neuroradiol 1985; 6: 767-772.

Matsuo A, Akiguchi I, Lee GC, McGeer EG, McGeer PL, Kimura J. Myelin degeneration in multiple system atrophy detected by unique antibodies. Am J Pathol. 1998; 153: 735-744.

Matsuo H, Takashima H, Kishikawa M, Kinoshita I, Mori M, Tsujihata M *et al.* Pure akinesia: an atypical manifestation of progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 1991; 54: 397-400.

Mattis S. Mental status examination for organic mental syndrome in the elderly patients. In: Bellack L, Karasu TB, editors. Geriatric psychiatry. New York: Grune & Stratton; 1973: 77-121.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K *et al.* A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos.Trans.R.Soc.Lond B Biol.Sci. 2001; 356: 1293-1322.

McCormick DA, Guyer PE, Thompson RF. Superior cerebellar peduncle lesions selectively abolish the ipsilateral classically conditioned nictitating membrane/eyelid response of the rabbit. Brain Res. 1982; 244: 347-350.

Middleton FA, Strick PL. A revised neuroanatomy of frontal subcortical circuits. In: Lichter DG, Cummings JL, editors. Frontal subcortical circuits in psychiatric and neurological disorders. Guildford, New York: 2000: 44-58.

Morris HR, Gibb G, Katzenschlager R, Wood NW, Hanger DP, Strand C *et al.* Pathological, clinical and genetic heterogeneity in progressive supranuclear palsy. Brain 2002; 125: 969-975.

Motoi Y, Takanashi M, Itaya M, Ikeda K, Mizuno Y, Mori H. Glial localization of four-repeat tau in atypical progressive supranuclear palsy. Neuropathology 2004; 24: 60-65.

Mueller EA, Moore MM, Kerr DC, Sexton G, Camicioli RM, Howieson DB *et al.* Brain volume preserved in healthy elderly through the eleventh decade. Neurology 1998; 51: 1555-1562.

Muqit MM, Mort D, Miskiel KA, Shakir RA. "Hot cross bun" sign in a patient with parkinsonism secondary to presumed vasculitis. J Neurol Neurosurg Psychiatry 2001; 71: 565-566.

Murata Y, Kawakami H, Yamaguchi S, Nishimura M, Kohriyama T, Ishizaki F *et al.* Characteristic magnetic resonance imaging findings in spinocerebellar ataxia 6. Arch.Neurol 1998a; 55: 1348-1352.

Murata Y, Yamaguchi S, Kawakami H, Imon Y, Maruyama H, Sakai T *et al.* Characteristic magnetic resonance imaging findings in Machado-Joseph disease. Arch.Neurol 1998b; 55: 33-37.

Murialdo A, Marchese R, Abbruzzese G, Tabaton M, Michelozzi G, Schiavoni S. Neurosyphilis presenting as progressive supranuclear palsy. Mov Disord. 2000; 15: 730-731.

Nakashima H, Terada S, Ishizu H, Tanabe Y, Yokota O, Ishihara T *et al.* An autopsied case of dementia with Lewy bodies with supranuclear gaze palsy. Neurol Res. 2003; 25: 533-537.

Nath U, Ben Shlomo Y, Thomson RG, Lees AJ, Burn DJ. Clinical features and natural history of progressive supranuclear palsy: a clinical cohort study. Neurology 2003; 60: 910-916.

Nath U, Ben Shlomo Y, Thomson RG, Morris HR, Wood NW, Lees AJ *et al.* The prevalence of progressive supranuclear palsy (Steele-Richardson- Olszewski syndrome) in the UK. Brain 2001; 124: 1438-1449.

Nelson HE. A modified card sorting test sensitive to frontal lobe defects. Cortex 1976; 12: 313-324.

Nelson H. The National Adult Reading Test (NART): Test Manual. Windsor: The NFER-NELSON Publishing Company Ltd.; 1982.

O'Sullivan M, Singhal S, Charlton R, Markus HS. Diffusion tensor imaging of thalamus correlates with cognition in CADASIL without dementia. Neurology 2004; 62: 702-707.

Oba H, Yagishita A, Terada H, Barkovich AJ, Kutomi K, Yamauchi T *et al.* New and reliable MRI diagnosis for progressive supranuclear palsy. Neurology 2005; 64: 2050-2055.

Ohshita T, Oka M, Imon Y, Yamaguchi S, Mimori Y, Nakamura S. Apparent diffusion coefficient measurements in progressive supranuclear palsy. Neuroradiology 2000; 42: 643-647.

Ohtsuka K, Igarashi Y, Chiba S. Cerebellar peduncle lesion without saccadic abnormalities. Ophthalmologica 1992; 204: 44-48.

Oka M, Katayama S, Imon Y, Ohshita T, Mimori Y, Nakamura S. Abnormal signals on proton density-weighted MRI of the superior cerebellar peduncle in progressive supranuclear palsy. Acta Neurol Scand. 2001; 104: 1-5.

Osaki Y, Ben Shlomo Y, Lees AJ, Daniel SE, Colosimo C, Wenning G *et al.* Accuracy of clinical diagnosis of progressive supranuclear palsy. Mov Disord 2004; 19: 181-189.

Ozawa T, Paviour D, Quinn NP, Josephs KA, Sangha H, Kilford L *et al.* The spectrum of pathological involvement of the striatonigral and olivopontocerebellar

systems in multiple system atrophy: clinicopathological correlations. Brain 2004; 127: 2657-2671.

Paling SM, Williams ED, Barber R, Burton EJ, Crum WR, Fox NC *et al.* The application of serial MRI analysis techniques to the study of cerebral atrophy in late-onset dementia. Med Image Anal. 2004; 8: 69-79.

Papp MI, Lantos PL. The distribution of oligodendroglial inclusions in multiple system atrophy and its relevance to clinical symptomatology. Brain 1994; 117 (Pt 2): 235-243.

Paviour DC, Lees AJ, Josephs KA, Ozawa T, Ganguly M, Strand C *et al.* Frontotemporal lobar degeneration with ubiquitin-only-immunoreactive neuronal changes: broadening the clinical picture to include progressive supranuclear palsy. Brain 2004a; 127: 2441-2451.

Paviour DC, Schott JM, Stevens JM, Revesz T, Holton JL, Rossor MN *et al.* Pathological substrate for regional distribution of increased atrophy rates in progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 2004b; 75: 1772-1775.

Pharr V, Uttl B, Stark M, Litvan I, Fantie B, Grafman J. Comparison of apraxia in corticobasal degeneration and progressive supranuclear palsy. Neurology 2001; 56: 957-963.

Pharr V, Litvan I, Brat DJ, Troncoso J, Reich SG, Stark M. Ideomotor apraxia in progressive supranuclear palsy: a case study. Mov Disord. 1999; 14: 162-166.

Pillon B, Gouider-Khouja N, Deweer B, Vidailhet M, Malapani C, Dubois B *et al.* Neuropsychological pattern of striatonigral degeneration: comparison with Parkinson's disease and progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 1995; 58: 174-179.

Pillon B, Deweer B, Michon A, Malapani C, Agid Y, Dubois B. Are explicit memory disorders of progressive supranuclear palsy related to damage to striatofrontal circuits? Comparison with Alzheimer's, Parkinson's, and Huntington's diseases. Neurology 1994; 44: 1264-1270.

Pirtosek Z, Jahanshahi M, Barrett G, Lees AJ. Attention and cognition in bradykinetic-rigid syndromes: an event-related potential study. Ann Neurol 2001; 50: 567-573.

Plummer DL. DispImage a Display and Analysis Tool for Medical Images. Revista Di Neuroradiologica 1992; 5: 489-495.

Reiss AL, Abrams MT, Singer HS, Ross JL, Denckla MB. Brain development, gender and IQ in children. A volumetric imaging study. Brain 1996; 119 (Pt 5): 1763-1774.

Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept.Mot.Skills 1958; 271-276.

Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. J Neurosci 2003; 23: 3295-3301.

Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ *et al.* One-year age changes in MRI brain volumes in older adults. Cereb.Cortex 2000; 10: 464-472.

Revesz T, Sangha H, Daniel SE. The nucleus raphe interpositus in the Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). Brain 1996; 119 (Pt 4): 1137-1143.

Revesz T, Daniel SE, Lees AJ, Will RG. A case of progressive subcortical gliosis associated with deposition of abnormal prion protein (PrP). J Neurol Neurosurg Psychiatry 1995; 58: 759-760.

Rey A. L'examen Clinique en Psychologie. Paris: PUF; 1958.

Richardson C, Steele JC, Olszewski J. Supranuclear Ophthalmoplegia, pseudobulbar palsy, nuchal dystonia and dementia. A clinical report on eight cases of heterogeneous system degeneration". Trans.Am.Neurol Assoc 1963; 88: 25-29.

Righini A, Antonini A, De Notaris R, Bianchini E, Meucci N, Sacilotto G *et al.* MR imaging of the superior profile of the midbrain: differential diagnosis between progressive supranuclear palsy and Parkinson disease. AJNR Am J Neuroradiol. 2004; 25: 927-932.

Robbins TW, James M, Owen AM, Lange KW, Lees AJ, Leigh PN *et al.* Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple system

atrophy in tests sensitive to frontal lobe dysfunction. J Neurol Neurosurg Psychiatry 1994; 57: 79-88.

Rojo A, Pernaute RS, Fontan A, Ruiz PG, Honnorat J, Lynch T *et al.* Clinical genetics of familial progressive supranuclear palsy. Brain 1999; 122 (Pt 7): 1233-1245.

Ros R, Gomez GP, Hirano M, Tai YF, Ampuero I, Vidal L *et al.* Genetic linkage of autosomal dominant progressive supranuclear palsy to 1q31.1. Ann Neurol 2005; 57: 634-641.

Rossor MN, Revesz T, Lantos PL, Warrington EK. Semantic dementia with ubiquitin-positive tau-negative inclusion bodies. Brain 2000; 123 (Pt 2): 267-276.

Royall DR, Mahurin RK, Gray KF. Bedside assessment of executive cognitive impairment: the executive interview. J Am.Geriatr Soc 1992; 40: 1221-1226.

Rusinek H, De Santi S, Frid D, Tsui WH, Tarshish CY, Convit A *et al.* Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. Radiology 2003; 229: 691-696.

Savoiardo M, Strada L, Girotti F, D'Incerti L, Sberna M, Soliveri P *et al.* MR imaging in progressive supranuclear palsy and Shy-Drager syndrome. J.Comput.Assist.Tomogr. 1989; 13: 555-560.

Savoiardo M, Girotti F, Strada L, Ciceri E. Magnetic resonance imaging in progressive supranuclear palsy and other parkinsonian disorders. J Neural Transm.Suppl 1994; 42: 93-110.

Scahill RI, Schott JM, Stevens JM, Rossor MN, Fox NC. Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. Proc.Natl.Acad.Sci.U.S.A 2002; 99: 4703-4707.

Schocke MF, Seppi K, Esterhammer R, Kremser C, Jaschke W, Poewe W *et al.* Diffusion-weighted MRI differentiates the Parkinson variant of multiple system atrophy from PD. Neurology 2002; 58: 575-580.

Schott JM, Simon JE, Fox NC, King AP, Khan MN, Cipolotti L *et al.* Delineating the sites and progression of in vivo atrophy in multiple system atrophy using fluid-registered MRI. Mov Disord. 2003; 18: 955-958.

Schrag A, Selai C, Quinn N, Hobart J. Measuring health-related quality of life in patients with progressive supranuclear palsy. Neurocase 2005; 11: 246-249.

Schrag A, Good CD, Miszkiel K, Morris HR, Mathias CJ, Lees AJ *et al.* Differentiation of atypical parkinsonian syndromes with routine MRI. Neurology 2000; 54: 697-702.

Schrag A, Ben Shlomo Y, Quinn NP. Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. Lancet 1999; 354: 1771-1775.

Schulz JB, Skalej M, Wedekind D, Luft AR, Abele M, Voigt K *et al.* Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. Ann.Neurol 1999; 45: 65-74.

Seo SW, Heo JH, Lee KY, Shin WC, Chang DI, Kim SM *et al.* Localization of Claude's syndrome. Neurology 2001; 57: 2304-2307.

Seppi K, Schocke MF, Wenning GK, Poewe W. How to diagnose MSA early: round table magnetic resonance imaging. J Neural Transm. 2005.

Seppi, K., Schocke, M. F., Mair, K. J., Jascke, W., Poewe, W.,.Wenning, G. K. Progressive degeneration on striatopallidal pathways in the Parkinson variant of multiple system atrophy: A longitudinal diffusion weighted MRI study. Mov Disord 19(Supplement 9), s334. 2004.
Ref Type: Abstract

Seppi K, Schocke MFH, Esterhammer R, Kremser C, Brenneis C, Mueller J *et al.* Diffusion-weighted imaging discriminates progressive supranuclear palsy from PD, but not from the parkinson variant of multiple system atrophy. Neurology 2003; 60: 922.

Silbert LC, Quinn JF, Moore MM, Corbridge E, Ball MJ, Murdoch G *et al.* Changes in premorbid brain volume predict Alzheimer's disease pathology. Neurology 2003; 61: 487-492.

Silbert PL, Gubbay SS, Khangure M. Multifocal astrocytoma masquerading as possible progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 1993; 56: 220-221.

Slachevsky A, Villalpando JM, Sarazin M, Hahn-Barma V, Pillon B, Dubois B. Frontal assessment battery and differential diagnosis of frontotemporal dementia and Alzheimer disease. Arch Neurol 2004; 61: 1104-1107.

Smith AD, Jobst KA. Use of structural imaging to study the progression of Alzheimer's disease. Br.Med Bull. 1996; 52: 575-586.

Smyth GE, Stern K. Tumours of the thalamus - a clinico-pathological study. Brain 1938; 61: 339-374.

Soliveri P, Rossi G, Monza D, Tagliavini F, Piacentini S, Albanese A *et al.* A Case of Dementia Parkinsonism Resembling Progressive Supranuclear Palsy Due to Mutation in the Tau Protein Gene. Arch Neurol 2003; 60: 1454-1456.

Soliveri P, Monza D, Paridi D, Carella F, Genitrini S, Testa D *et al.* Neuropsychological follow up in patients with Parkinson's disease, striatonigral degeneration-type multisystem atrophy, and progressive supranuclear palsy. J Neurol Neurosurg Psychiatry 2000; 69: 313-318.

Specht K, Minnerop M, Abele M, Reul J, Wullner U, Klockgether T. In vivo voxel-based morphometry in multiple system atrophy of the cerebellar type. Arch.Neurol 2003; 60: 1431-1435.

Steele JC, Richardson JC, Olszewski J. Progressive Supranuclear Palsy. Arch Neurol 1964; 10: 333-359.

Stell R, Bronstein AM. Eye movement abnormalities in extrapyramidal diseases. In: Marsden CD, Fahn S, editors. Movement Disorders 3. London: Butterworth-Heinemann; 1994: 88-113.

Stern MB, Braffman BH, Skolnick BE, Hurtig HI, Grossman RI. Magnetic resonance imaging in Parkinson's disease and parkinsonian syndromes. Neurology 1989; 39: 1524-1526.

Stieltjes B, Kaufmann WE, van Zijl PC, Fredericksen K, Pearlson GD, Solaiyappan M *et al.* Diffusion tensor imaging and axonal tracking in the human brainstem. NeuroImage 2001; 14: 723-735.

Symms M, Jager HR, Schmierer K, Yousry TA. A review of structural magnetic resonance neuroimaging. J Neurol Neurosurg Psychiatry 2004; 75: 1235-1244.

Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? JAMA 1999; 282: 790-795.

Tison F, Yekhlef F, Chrysostome V, Balestre E, Quinn NP, Poewe W *et al.* Parkinsonism in multiple system atrophy: natural history, severity (UPDRS-III), and disability assessment compared with Parkinson's disease. Mov Disord 2002; 17: 701-709.

Tokumaru AM, Kamakura K, Maki T, Murayama S, Sakata I, Kaji T *et al.* Magnetic resonance imaging findings of Machado-Joseph disease: histopathologic correlation. J Comput.Assist.Tomogr. 2003; 27: 241-248.

Tolosa E, Valldeoriola F, Marti MJ. Clinical diagnosis and diagnostic criteria of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome). J Neural Transm.Suppl 1994; 42: 15-31.

Tsuboi Y, Slowinski J, Josephs KA, Honer WG, Wszolek ZK, Dickson DW. Atrophy of superior cerebellar peduncle in progressive supranuclear palsy. Neurology 2003; 60: 1766.

Verny M, Duyckaerts C, Agid Y, Hauw JJ. The significance of cortical pathology in progressive supranuclear palsy. Clinico-pathological data in 10 cases. Brain 1996a; 119 (Pt 4): 1123-1136.

Verny M, Jellinger KA, Hauw JJ, Bancher C, Litvan I, Agid Y. Progressive supranuclear palsy: a clinicopathological study of 21 cases. Acta Neuropathol.(Berl) 1996b; 91: 427-431.

Voneida TJ. The effect of brachium conjunctivum transection on a conditioned limb response in the cat. Behav.Brain Res. 2000; 109: 167-175.

Wakabayashi K, Ikeuchi T, Ishikawa A, Takahashi H. Multiple system atrophy with severe involvement of the motor cortical areas and cerebral white matter. J Neurol Sci. 1998; 156: 114-117.

Warmuth-Metz M, Naumann M, Csoti I, Solymosi L. Measurement of the midbrain diameter on routine magnetic resonance imaging: a simple and accurate method of differentiating between Parkinson disease and progressive supranuclear palsy. Arch.Neurol 2001; 58: 1076-1079.

Warrington EK. Manual for the Recognition Memory Test for words and faces. Windsor: The NFER-NELSON Publishing Company Ltd.; 1984.

Watanabe H, Saito Y, Terao S, Ando T, Kachi T, Mukai E *et al.* Progression and prognosis in multiple system atrophy: an analysis of 230 Japanese patients. Brain 2002; 125: 1070-1083.

Weeks RA, Scaravilli F, Lees AJ, Carroll C, Husain M, Rudge P. Cerebral amyloid angiopathy and motor neurone disease presenting with a progressive supranuclear palsy-like syndrome. Mov Disord. 2003; 18: 331-336.

Weschler D. Manual for the Wechsler Adult Intelligence Scale - Revised. New York Psychological Corp; 1981.

Whitwell JL, Crum WR, Watt HC, Fox NC. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. AJNR Am.J.Neuroradiol. 2001; 22: 1483-1489.

Whone AL, Watts RL, Stoessl AJ, Davis M, Reske S, Nahmias C *et al.* Slower progression of Parkinson's disease with ropinirole versus levodopa: The REAL-PET study. Ann Neurol 2003; 54: 93-101.

Will RG, Lees AJ, Gibb W, Barnard RO. A case of progressive subcortical gliosis presenting clinically as Steele-Richardson-Olszewski syndrome. J Neurol Neurosurg Psychiatry 1988; 51: 1224-1227.

Williams DR, de Silva R, Paviour DC, Pittman A, Watt HC, Kilford L *et al.* Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. Brain 2005; 128: 1247-1258.

Winikates J, Jankovic J. Vascular progressive supranuclear palsy. J Neural Transm.Suppl 1994; 42: 189-201.

Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC. Automated image registration: I. General methods and intrasubject, intramodality validation. J Comput.Assist.Tomogr. 1998; 22: 139-152.

Yagishita A, Oda M. Progressive supranuclear palsy: MRI and pathological findings. Neuroradiology 1996; 38 Suppl 1: S60-S66.

Yamauchi H, Fukuyama H, Nagahama Y, Katsumi Y, Dong Y, Konishi J *et al.* Atrophy of the corpus callosum, cognitive impairment, and cortical hypometabolism in progressive supranuclear palsy. Ann.Neurol. 1997; 41: 606-614.

dementia, progressive supranuclear palsy, and Alzheimer's disease. J.Neurol.Neurosurg.Psychiatry 2000; 69: 623-629.

Yamauchi H, Fukuyama H, Nagahama Y, Katsumi Y, Hayashi T, Oyanagi C *et al.* Comparison of the pattern of atrophy of the corpus callosum in frontotemporal dementia, progressive supranuclear palsy, and Alzheimer's disease